1994-09

Design of a predictive recruiter success model (PRiSM)

Hernandez, Alejandro S.

Monterey, California. Naval Postgraduate School

NAVAL POSTGRADUATE SCHOOL
Monterey, California

THESIS

DESIGN OF A PREDICTIVE RECRUITER SUCCESS MODEL (PRiSM)

by

Alejandro S. Hernandez

September, 1994

Thesis Advisors: James G. Taylor
Ronald A. Weitzman

Approved for public release; distribution is unlimited.

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1 AGENCY USE ONLY *(Leave blank)*	2 REPORT DATE September 1994	3. REPORT TYPE AND DATES COVERED Master's Thesis

4. TITLE AND SUBTITLE: Design of A Predictive Recruiter Success Model (PRiSM)	5.
6. AUTHOR(S) Alejandro S. Hernandez	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey CA 93943-5000	8.
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) US Army Recruiting Command, PA&E Division, Fort Knox, KY	10. USAREC

11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

12a. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution is unlimited.	12b. DISTRIBUTION CODE *A

13. ABSTRACT *(maximum 200 words)*

This thesis describes the formulation and validation of a multiple linear regression model that predicts recruiter success rates. The model's primary purpose is to improve the recruiter selection process by helping to reduce recruiter reliefs. Using recorded information on over 400 members of two active-duty recruiting battalions together with the results of an administered sales ability test,a database was constructed for use in regression analyses. Recruiter success was defined as the response variable in specific, quantifiable terms. Potential predictive variables were identified to reflect the ideal traits of a successful recruiter. The method of Mallow's Coefficient C_p in conjunction with hypothesis tests, was used to develop the final predictive model. Residual analyses, data-splitting, and cross-validation methods assured the appropriateness and adequacy of the final model to describe and predict recruiter success. However, this model is limited by the fact that all sales ability data was collected using the " present-employee " method. For the purpose of calculating potential cost savings, an analysis using the Taylor and Russell tables was conducted. Cost savings expected from use of the model amounted to nearly $3.38 million annually.

14. SUBJECT TERMS US Army Recruiting Command (USAREC), Predictive Recruiter Success Model (PRiSM), Sales Comprehension Test (SCT), Delayed Entry Program (DEP), Primary Military Occupational Skill (PMOS), Mallow's Coefficient	15. NUMBER OF PAGES 202
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT UL

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18

Approved for public release; distribution is unlimited.

DESIGN OF A PREDICTIVE RECRUITER SUCCESS MODEL (PRiSM)

Alejandro S. Hernandez

Captain, United States Army

B.S. Civil Engineering, ... States Military Academy, 1985

Submitted in partial fulfillment

of the requirements for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the

NAVAL POSTGRADUATE SCHOOL

September 1994

Author: [redacted]

Alejandro S. Hernandez

Approved by: [redacted]

James G. Taylor, Thesis Advisor

[redacted]

Ronald A. Weitzman, Thesis Advisor

[redacted]

George C. Prueitt, Second Reader

[redacted]

Peter Purdue, Chairman, Department of Operations Research

ABSTRACT

This thesis describes the formulation and validation of a multiple linear regression model that predicts recruiter success rates. The model's primary purpose is to improve the recruiter selection process by helping to reduce recruiter reliefs. Using recorded information on over 400 members of two active-duty recruiting battalions together with the results of an administered sales ability test, a database was constructed for use in regression analyses.

Recruiter success was defined as the response variable in specific, quantifiable terms. Potential predictive variables were identified to reflect the ideal traits of a successful recruiter. The method of Mallow's Coefficient C_p, in conjunction with hypothesis tests, was used to develop the final predictive model. Residual analyses, data-splitting, and cross-validation methods assured the appropriateness and adequacy of the final model to describe and predict recruiter success. However, this model is limited by the fact that all sales ability data was collected using the " present-employee " method. For the purpose of calculating potential cost savings, an analysis using the Taylor and Russell tables was conducted. Cost savings expected from use of the model amounted to nearly $3.38 million annually.

TABLE OF CONTENTS

ACKNOWLEDGMENTS

I wish to express my sincere gratitude to Professor Bard Mansager, a member of the Recruiter Selection Support system, for laying the groundwork for this study.

The data used in this thesis was collected and shared with CPT(P) Todd A. Buchs who did a related study. His positive approach to the problem of data collection and database construction advanced both studies immensely.

CPT Tom Nelson was invaluable in condensing the original database information received from USAREC into a coherent form. His knowledge of database manipulation saved a great deal of valuable time.

I also acknowledge two members of the Baltimore Recruiting Battalion. Throughout this study, CPT(P) George Getczy, Battalion Operations Officer, expertly and energetically supported the project with his insight. Battalion Sergeant Major Joseph Quig, a self-appointed mentor, whose vast knowledge of recruiters, recruiting, and leadership, was invaluable in providing a realistic glimpse into the recruiting arena.

Mr. Robert Hamilton, Analyst for the Defense Manpower Data Center, supported my efforts with timely and accurate information.

Lastly, I acknowledge the soldiers of the Santa Anna and Baltimore Recruiting Battalions for their support and cooperation in making this study possible. Their professionalism and personal sacrifices continue to serve and strengthen the U.S. Army.

EXECUTIVE SUMMARY

The United States Army Recruiting Command (USAREC) needs top-quality young men and women to meet the recruiting needs of the Modern Army. The field recruiter is USAREC's front-line representative in this mission. The annual number of recruiter failures (caused by the recruiter's inability to meet mission quotas) has been, and continues to be, at intolerably high levels. Data for the past four years reveals that each year recruiter reliefs have varied between seven and ten percent in a total population of 6700 recruiters. The percentage of reliefs is much greater during the first nine months of recruiting duty. These reliefs cost USAREC nearly five million dollars a year, and erode unit efficiency and morale. Therefore, it is imperative to determine if an "instrument" can be developed to improve USAREC's recruiter selection process, thereby reducing the number of recruiters relieved, and saving millions of dollars yearly while increasing USAREC's productivity.

Although recruiters are selected from the top ten percent of their respective primary branches, their success in recruiting is not guaranteed. Soldiers who have had successful careers in their primary branch still have difficulty succeeding in recruiting. The rigors of the recruiting environment and the change in mission tasks demand individuals who have an aptitude for this specific military occupation. To measure an individual's potential success as a recruiter, it was determined that an instrument to distinguish successful from unsuccessful recruiter traiits, had to be developed.

This study used multiple regression analysis techniques to develop and validate a model that can predict a recruiter candidate's potential for success. A database was constructed from the recruiting records of over 400 field recruiters from two different active-duty recruiting battalions. Key to the construction was the administration of a test that measured a recruiter's sales aptitude, a measure

absent from all recruiter records, but one that is considered to be an important trait of a successful recruiter. A measure of effectiveness (MOE) for recruiter success was developed from the detailed history of assigned and achieved missions of each recruiter in the sample. A common time interval (the first nine months of recruiting duty) was selected to calculate recruiter success, to ensure that each sample was measured on an equal basis. The MOE also incorporated the effects of the Army's Delayed Entry Program (DEP) and the different types of recruiting missions that could be assigned. Guided by USAREC's description of a successful recruiter, potential predictive variables were identified from information in the database. Using the MOE as the dependent variable and key recruiter traits as independent variables, regression methods were employed to develop a predictive model.

Mallow's Coefficient, C_p, was used for variable selection and to develop the initial multi-linear regression model. Mallow's C_p was instrumental in selecting the variables because of its value in choosing the variables that best describe the data. Hypothesis tests on the full and reduced models, combined with hypothesis tests on each estimated coefficient, produced a final predictive model with a correlation coefficient of 0.3082. The final model contained four variables that measure certain characteristics of a recruiter candidate: the Armed Forces Qualification Test (AFQT) score, Sales Comprehension Test (SCT) score, Gender, and Primary Military Occupational Skill (PMOS), the last two variables being binary.

Residual analysis revealed that a linear regression model was appropriate for describing the dependence of the MOE on these four independent variables. To validate the model's predictive ability, data splitting and cross-validation methods were used. These methods indicated that the model's predictive ability was well within its expected limitations.

The annual TTE failure rate for a recent year was applied in a cost-benefit analysis using the Taylor and Russell tables, and an estimated number of failures among selectees was computed. Potential savings for USAREC were calculated by comparing the actual number with the expected number of failures using the screening capabilities of the model. The amount of potential savings was substantial, amounting to nearly $3.38 million annually.

The development and validation of the final model indicated that an instrument to aid in improving the current recruiter selection process is feasible and promising. The benefits of the model are dependent on the time period and manner of application. Currently, the only additional information required to employ the model is a measure of sales ability. Collection of data and application of the model can be done concurrently. Since recruiters must be Noncommissioned Officers (NCO), one opportunity for testing exists during the soldier's attendance at the Primary Leadership Development Course (PLDC), a required school for any future NCO. There may be other alternatives, but evaluation of these options is not within the scope of this study. The model's prediction can be recorded in a soldier's military files, for use by assignment officers and members of the Recruit the Recruiter Program.

One shortcoming in the development of the model was the use of the "present-employee" method for collecting sales ability data. Because this procedure was used, the full spectrum of recruiter data was not obtained. Lack of information for failed recruiters was most notable and caused homogeneity in the sample. This homogeneity suggests prudence when employing the model, and indicates that further research is needed. Other applications of the model, as well as recommendations to improve it, are outlined in this study.

I. INTRODUCTION

The primary mission of the United States Army is to deter war by being prepared to fight and win on the battlefield. The United States Army Recruiting Command (USAREC) is responsible for supporting this critical mission by providing quality young men and women to meet the needs of the Modern Army. USAREC accomplishes this mission through the use of the field recruiter. It is the recruiter who must meet the stringent quotas set by Congress for each of the nation's military services. These quotas are based on the needs of national security objectives. In this light, the ability of USAREC to accomplish its mission remains critical to national defense.

A. GENERAL BACKGROUND

U.S. Army recruiting success has been declining in recent years. As the service with the largest number of personnel, the Army should receive approximately forty-two percent of all service enlistments. In FY93, the Army claimed a mere twenty-eight percent. Various explanations have been offered for the shortfall, but USAREC considers recruiter performance as foremost among these.

Moreover, annual recruiter losses have been intolerable. In the past four years, the proportion of recruiters who have been relieved of duty has varied from seven to ten percent in a population of 6,700 recruiters. These reliefs have been costly in the forms of monetary losses, unit inefficiency, and low morale. Over seventy-one percent of these reliefs have resulted from recruiter ineffectiveness: the recruiter's inability to produce an assigned enlistment quota for a given period. This policy to relieve ineffective recruiters has been institutionalized in Army Regulations (AR) 601-1, which sets policies for USAREC recruiters. These recruiters' poor performances contribute directly to failure in USAREC's primary mission and indirectly to degraded mission performance of the U.S. Army as a whole.

Although recruiters are selected from the top ten percent of their respective primary branches, their success in recruiting is not guaranteed[1]. Soldiers who have successful careers in their primary branch still have difficulty succeeding in recruiting. As a result, these otherwise successful soldiers receive unfavorable efficiency reports, or are released from recruiting duty before an efficiency report is required. In the latter case, the soldier is still labeled as being a soldier who was not able to satisfactorily perform an assigned duty. To prevent these negative outcomes and improve the command's productivity it would be extremely convenient to be able to determine which soldiers should, and should not, be assigned to recruiting duty.

A USAREC-sponsored research project at the Naval Postgraduate School (NPS) found that, other than an interview from a recruiter's commander, or possibly a basic screening from the Recruit the Recruiter Program (RRP) team, a candidate recruiter did not undergo any type of formal screening test, although such a policy is followed for all other jobs in the U.S. Army. Thus, recruiting is the only Military Occupational Skill (MOS) in the U.S. Army which does not have this requirement. AR 601-1's selection criteria focus on *administrative* deficiencies that would prevent a soldier from becoming a recruiter (overweight soldiers, or those with marital or financial problems are the target of the policies). The source of this problem is that currently, there are no screening tests available that directly measure the skills which a successful recruiter needs. Further, salesmanship ability, a key trait of a successful recruiter, is not a skill tested at any level in the U.S. Army.

---

[1]Major Alan Poikenen from USAREC PA&E, during an in progress report (IPR), April 1994, stated that all recruiters are in the top ten percent of their branch

USAREC's goal is to improve the recruiter selection process. An instrument that can screen undesirable recruiter candidates from recruiting duty will help meet USAREC's goal of increasing recruiter production and decreasing recruiter failure. With a screening tool that reduces the number of unsuccessful recruiters, USAREC will be able to reduce the number of recruiters relieved from duty and increase the productivity of the command.

B. STUDY BACKGROUND

In June 1993, USAREC requested NPS to establish a study group to develop a tool to help in selecting potentially successful recruiters. Subsequently, an interdisciplinary team, with experts in systems analysis, behavioral testing, and applied statistics, was formed. Since no previous study had been conducted in this area of recruitment, it was left to the study group to find a methodology that would help USAREC in selecting recruiters. Intermediate goals were to define the traits common to sucessful recruiters, find quantifiable measures for these traits, and to clearly define recruiter success, so that it could be used as a measure of effectiveness (MOE) for recruiter performance.

The absence of a measure for recruiter sales skills prompted the team to search for a test that would measure this quality. The test would have to meet the criteria of being easy to administer and yielding results that could logically be interpreted to measure salesmanship ability. Several tests were considered, but most were inadequate for the study's purpose. The Sales Aptitude Test by the Employers' Test & Services Associates took sixty to seventy minutes to administer and involved complicated instructions to focus on 31 items as part of a *series* of tests. The Sales Aptitude Check List by the Science Research Associates applied only to people who have had sales experience. The Sales Motivation Inventory, with 75 separate items, appeared too long and time-consuming. [Ref. 1, p. 1]

The most promising of these tests was felt to be the Sales Comprehension Test (SCT). The SCT was developed by Dr. Martin M. Bruce to measure an individual's comprehension and appreciation of basic salesmanship skills. Previous studies have validated the test's results with specific groups such as salesmen, sales students, factory workers, and other sales and non-sales groups. [Ref. 2, pp. 3-6] Discussions with instructors from the Army Recruiting School (ARS) and recruiting battalion leaders indicated that the SCT emphasizes many of the skills that the ARS teaches new recruiters. A separate study has been launched to validate the SCT for recruiter selection in the Army[2].

The other issue of concern was the definition of success. USAREC manuals refer to a single standard, but subordinate units give this standard different operational definitions. Since each recruiter is evaluated on his production seventeen times every year, a single measure of success that captures the true performance of a recruiter must be used. Twelve of these measures are monthly quotas while the other five are aggregates in the form of quarterly and yearly quotas.

A database was created from USAREC's data collection system, the Army Research Institute (ARI), Defense Manpower Data Center (DMDC), surveys, and administered tests. The database includes the data of two chosen battalions of active recruiters and one Basic Noncommissioned Officer Course (BNCOC) class. The two active duty recruiting battalions were chosen on the basis of time, resources, and opportunity. The configuration of the database and an underlying need to select recruiters based on predicted performances suggested a multiple linear regression approach. The resulting model would use a selected measure of success for each recruiter as the dependent variable and individual recruiter traits as possible independent variables.

---

[2]A separate thesis that studies the validity of the Sales Comprehension Test as a tool to predict recruiter success was pursued by CPT Todd Buchs. The goal of the study was to validate the SCT's predictive ability in the recruiter population.

C. LIMITATIONS AND ASSUMPTIONS

Time is always of the essence. Inefficient recruiters cause USAREC to lose potential enlistments and to overspend funds for nonproductive reasons. Since it would require a minimum of nine months to track the success, or failure of a sample group of recruiters, the study used recruiters already in the field. Their characteristic information was extracted from the USAREC database. Recruiting training experience, familiarity with a sales environment, and other factors may bias the data obtained from this group. Because the study must view the recruiters as soldiers who may become recruiters, it must assume that their attributes before they were recruiters have stayed constant. These traits can possibly be used as factors to indicate future performance.

Because no recruiter had a measure of sales ability recorded, the SCT had to be administered to the recruiters in the study. The availability of funds, opportunity, and time to administer the test were resources the study group did not have in abundance. Two hundred seventy-six soldiers were tested. Since the data for a single recruiter would not be complete without a measure of his sales skills, the number of tested recruiters limited the number of recruiters used in the database.

Administration of a predictive test to an individual **after** that individual has been employed long enough to have a criterion measure available is called the "present-employee" method. An alternative method of validity testing is the "follow-up" method that involves administering a test **before** an individual is employed and comparing the criterion measure after enough data has been collected to compute the measure. [Ref. 3, pp. 114-115] Because the present-employee method of administering the SCT was used, the recruiters sampled were assumed to be homogeneous. Since all the recruiters being tested were still in the command, it was assumed that the command considered them "successful". The test would not be able to measure the sales ability of recruiters who were considered failures. This homogeneity in the sampled group may negatively affect the correlation of the independent variables with the dependent variable in the study.

Data in this study were not collected in a time sequence. Time sequenced data collection often leads to dependent measures. The MOE that is described in Chapter II was obtained at a predetermined number of months for every recruiter in the study. Recruiters, not time periods, were the units of measurement. The error terms are thus reasonably assumed to be independent.

D. PROBLEM STATEMENT

Can a mathematical model that explains and predicts variations in recruiter success be formulated, and, if so, what is its value as a predictive instrument of recruiter performance?

E. APPROACH TO MODEL DEVELOPMENT

The approach to this problem is a traditional least-squares multiple linear regression. The first steps in formulating the model is to determine the proper definition of the measure of effectiveness (MOE) and to identify the key variables that predict this measure. Using these as dependent and independent variables, respectively, the model can act as a predictive tool to aid USAREC in selecting recruiters with increased potential for success and less probability of failure. A set of characteristic values, which describe a recruiter candidate, can be entered as independent variables into the formula to predict potential success.

USAREC Manual 100-5 identifies the traits of a successful recruiter to be: (1) Salesmanship Ability, (2) Energy and Enthusiasm, (3) Communication Skills, and (4) Planning and Organizing Skills. USAREC does not propose any quantifiable measures for these traits. To determine if these qualities have a significant relationship with recruiter success, tests that measure these traits must be found and administered to recruiters. Each test must be evaluated to ensure that it is a valid measure to interpret the resulting model with more certainty. Other factors will also be investigated for use as measures or indicators of these key attributes.

Other characteristics of a recruiter may also influence his success. A simple regression on each of these possible variables may show a correlation between it and success. A non-zero correlation between the variable, or a function of the variable, and the MOE *may* suggest its inclusion in the model. A statistical test of the variable's coefficient in the multiple regression model will show if the variable has a significant effect on success. Covariance between independent variables will be examined and dealt with to make the model more robust and its results easier to interpret. Data-splitting and cross-validation methods will be used to validate the final model.

F. ESTIMATION OF USAREC SAVINGS

An estimate of cost savings to USAREC will be conducted in this study. A hypothetical recruiting class will be used to compare the losses from the current selection procedures and the savings in using the formulated model. The model can be used to identify those who should not have been selected for recruiting. Using dollar figures for the training cost of one recruiter, a cost for a group of failures will be calculated. The model will be used to compute a predicted MOE for each hypothetical recruiter. Varying the selection criteria in reference to the predicted MOE that USAREC could use, different amounts of savings can be computed by screening out "recruiters" who would have failed before funds were expended on them. These figures will give a general idea of the cost savings the application of the model can generate for USAREC.

II. LITERATURE REVIEW AND VARIABLE FRAMEWORK

Past studies concerning recruiters have focused on enlistees (the product of recruiter success), and not on the recruiters themselves. The study at hand focuses on the recruiter himself and, as a result, does not have many resources to use as a pattern for the study Examination of some approaches that have been used in previous recruiting studies can suggest ideas. The majority of this chapter specifies the reasons certain factors are considered for inclusion in the predictor model. Preliminary analysis by statistical and graphical methods will be presented to justify selection or non-selection of each variable.

A. LITERATURE REVIEW

1. Regression: A Standard Approach to Prediction

Two recent studies done in military recruiting helped in choosing a methodology to develop a mathematical model that predicts recruiter performance. Although neither directly studied recruiters, each dealt with predicting results. Research done in the Delayed Entry Program (DEP) concentrated on the probability that an enlistee would report for active duty, based on the time he spent in the DEP. [Ref. 4, p. 1] That study reported that the largest determinants of DEP loss were the *personal characteristics* of the enlistee. The study developed a multiple regression model that included factors such as the enlistee's age, gender, and race, to predict the probability that he would be a DEP loss. With these characteristics (along with other factors) as explanatory variables and the probability that the enlistee would report as the response variable, a regression model was developed. The other recruiting study dealt with the Soldier Retention Bonus (SRB). The SRB analysis was aimed at predicting the average length of service for a soldier who received an SRB. [Ref. 5, p. 14] In this case, the explanatory variables were the terms of the SRB and the response variable was the soldier's length of service in the U.S.

Army after receiving the SRB. Again, a regression model was generated from a database containing a history of past SRB contracts and the individual's service years. The SRB study focused on the soldier's characteristics, such as Military Occupational Skill, age, and other traits to predict the number of months that he would remain in active duty after receiving the bonus.

These two studies are reported in recent theses by NPS students. The framework of the recruiter selection problem is very similar. The focus is prediction. Various characteristics of the recruiter may indicate a propensity for success or failure. Unlike the previous two studies, this study does not enjoy a wealth of useful information on recruiters for analysis. Much of the data collected on aptitude and performance before a soldier becomes a recruiter are subjective evaluations. However, prediction is still the key. Regression analysis may still be conducted if a suitable database can be created from reliable sources.

2. Historical Criteria for Predictive Models

Correlation between predictive variables and actual performance measures have been studied at great length. Information from these studies are the benchmark which this study uses as a goal. A study by Garrett found that a correlation of 0.40 is the correlation subgroup that indicates "...reasonable and probably significant correlation..." [Ref. 6, p. 52] Further studies by Ghiselli reports that the average correlation between selective screening tests and job proficiency (actual job performance) was 0.20. [Ref. 7, p. 357] Later studies by Taylor and Russell validated Ghiselli's reports and described the use of predictive correlations, combined with selection ratios and success proportions, to derive a desired proportion of successful selectees. [Ref. 7, p. 361]

3. Further Readings

Since the factors which may prove to be useful predictive variables are unclear, qualitative and categorical traits are also considered. These traits represent "Non-Intellective General Factors" (NIGF). A study by Spearman notes that these traits govern an individual's application of his given abilities through intangible measures of determination, will, drive, etc. [Ref. 6, p. 131] Recruiters are in an environment where these characteristics are essential for their survival in the trade.

B. DATA COLLECTION AND DATABASE CONSTRUCTION

Outside agencies were instrumental in collecting information. As previously discussed, information on active recruiters would be used to construct the database. Because it was unknown which factors would truly influence success, model specification error could occur. It was felt that as much information that could logically be linked with success should be collected. Each variable could be scrutinized more closely at later stages of the study.

Much of the information came from USAREC's data system. It provided such specifics as age, sex, length of service, primary military occupational specialty (PMOS), missions achieved, and other statistics. DMDC and ARI were able to provide test results for the Armed Services Vocational Aptitude Battery (ASVAB) and the Armed Forces Qualification Test (AFQT). Other information, such as how a recruiter was assigned to recruiting or if a detail recruiter planned to change his PMOS to recruiting, were obtained by administering a questionnaire. A vital part of data collection was obtaining a measure of sales ability; data not available from recruiter records or any other source.

This absence of information on sales ability posed two critical questions. These questions were critical because the answers would determine the size of the database. The first was **whom** to test. The second was how **many** to test. Each question was constrained by opportunity, time, and funding.

A primary concern of the study was that the sample be as representative of the Army population as possible. USAREC consists of four brigades, comprised of forty recruiting battalions. USAREC uses the SMART BOOK system to rank order battalions, based on their production. [Ref. 8, pp. 2-4] The goal of the study was to administer the SCT to three battalions and collect data on approximately 400 recruiters. One battalion from each of the upper, middle, and lower third of the SMART BOOK rankings would be chosen.

This process of Unequal Probability of Selection (UPS) lent itself naturally to the population under investigation. The population is administratively divided into blocks, or units. The stratified nature of the population also made it necessary to examine the possibility of significantly different variances among strata. The UPS sampling method was coupled with small-scale sampling, involving limited resources that could affect the extent of the generalizations of the findings and their degree of accuracy. [Ref. 9, p. 103]

Opportunity and funding were key to determining the number of recruiters to whom the SCT could be administered. During the data collection stage of the study, only two military units could be visited. The Baltimore Recruiting Battalion and the Santa Anna Recruiting Battalion, ranked in the upper and middle third respectively, took the SCT. Only Regular Army (RA), active recruiters were tested. The testing produced results from 276 recruiters. The initial size of the database began with the records of these 276 individuals (Appendix A). Combining the two units into one database raised concerns about the possible effects of different unit variances and means for the MOE that will be described in the next section. An F-test on unit variances and a t-test on unit means were conducted to evaluate the possibility of significant differences between the two units. The test results in Appendix B indicate that there are no significant differences between the two units in terms of means and variances.

C. VARIABLE DESCRIPTIONS

1. Success MOE

USAREC Manual 100-5 specifically states that a recruiter who meets or exceeds a pre-established quota of enlistments is successful. The general spirit of this policy is clear. What is unclear is the time period to which the policy refers. A recruiter must make quota seventeen times each year: monthly, quarterly, and annually. These time periods are known throughout the recruiting community as the seventeen races[3]. The question is, on which time period should the study focus?

USAREC Regulations and Manuals and AR 601-1 allude to monthly quotas as the critical mission for the recruiter. The USAREC Manual emphasizes the repercussions of what a single missed enlistment for each recruiter in any given month would mean to the Army. Likewise, USAREC Pamphlet 350-11 directs recruiters who demonstrate deficiencies in **monthly** Individual Sustainment Training (IST), or are nonproductive, be enrolled in the station commander's Individual Training Plan (ITP). AR 601-1 further stipulates that **monthly** reevaluation is required for recruiters who display difficulty in meeting mission requirements.

This study will therefore define success as meeting monthly missions. Quarterly and annual quotas are based on these monthly missions. As the USAREC Manual explains, the damage to unit readiness in Army units is most profound when an anticipated replacement is not received in the same month the need arises. Increasing production the following month does little to compensate a unit that has missed the replacement the previous month, because new replacements are also needed for the following months. The unit has now fallen behind in their acquisition of replacements. This situation is echoed for other units which have not received needed replacements. These conditions are not affected nearly as much by quarterly and annual quotas,

---

[3]In a background interview with Sergeant Major (SGM) Joseph B. Quig III, SGM for Baltimore Recruiting Battalion, Maryland, the subject of the "seventeen races" was discussed. SGM Quig embellished the hardships which recruiters must undergo to meet twelve months, four quarters, and one year's worth of quotas to satisfy the needs of the Army and the demands of the recruiting battalion.

which may show high percentages of the number of recruitments, but fail to show the detailed negative effects of missed monthly missions.

Recruiters face the additional problem of DEP Loss. The DEP was instituted in 1960 to help U.S. Army Personnel Command (PERSCOM) schedule the entrance of new soldiers into active duty. [Ref. 10, p. 253] An enlistee is allowed a maximum of twelve months to delay his or her induction into the Army after signing a contract. However, the enlistee can choose not to report to active duty without any major adverse actions. Studies have shown that the longer an enlistee is in the DEP, the greater the probability that he or she will not report. [Ref. 10, p. 265]

The recruiter's responsibility is to ensure that an enlistee in the DEP reports for active duty. Accomplishment of this task requires continued contact with the enlistee, along with scheduled activities that will maintain the enlistee's interest in his or her choice. An enlistee who fails to report for active duty is counted against the recruiter who signed him or her to a contract. The recruiter must compensate for this loss in the current month, in addition to the *mission* already assigned. This study will count DEP Loss against the recruiter's total mission achievements for the month he or she signed the enlistee into the DEP.

The MOE for measuring recruiter success will be the proportion of months a recruiter has made mission. The USAREC database provides monthly statistics for the number of missions a recruiter was assigned and the number achieved, taking DEP Loss into account. The success MOE for each recruiter is calculated by granting **one** point for every month that a recruiter met or exceeded his mission and **zero** for each month he missed mission. A recruiter meets mission by enlisting the number of prospects he has been assigned for that month. Enlistments above the given quota are irrelevant to how that month is scored for the MOE. If a recruiter is not given a mission for a given month, then that month is automatically scored one point. However, this situation very rarely occurs. Less than 2% of the recruiters in the sample have ever experienced a month in which he was given no mission. This result is understandable: USAREC would be wasting an asset by not employing a soldier for the task he has been trained to do. Totaling these points and dividing by the number of months he was a recruiter resulted in the proportion of months the recruiter made mission.

Recruiters are given two general types of missions: Graduate Senior Category A (GSA), which include premium enlistees, and other-than-premium missions, which will be referred to as volume (VOL) missions in this study. On the average, Department of the Army (DA) policy apportions missions as sixty-seven percent GSA and the remaining fraction as VOL[4]. To improve the MOE as a measure of success, separate calculations were made for the percentage of months a recruiter made GSA missions and the percentage of months he or she made VOL missions. Each percentage was then weighted in accordance with DA policy fractions, 0.67 and 0.33, respectively. The sum produced a weighted percentage that defined a recruiter's overall success.

A similar method for calculating an MOE for recruiter success would take the sum of the weighted cumulative percentages for GSA and VOL. However, this method fails to capture the true performance of a recruiter. Since each month is regarded as the critical time period, a cumulative approach would taint the true picture. A recruiter who fails to make mission one month and exceeds his mission the next does not have the excess tacked to the previous month's total. The damage to the U.S. Army mission has already been done. As a result, the previous month remains a month in which he failed and the current month is a month in which he succeeded. A cumulative approach could very

[4]CPT(P) George Gezcy. Operations Officer, Baltimore Recruiting Battalion, reported these fractions of GSA and other than GSA accession goals as unwritten, but standard policy for recruiting. Later conversations with MAJ Alan Poikonen, USAREC PA&E, confirmed these figures.

possibly show that a recruiter made ninety percent of all his missions, but only accomplished his mission on fifty percent of the months.

The disadvantages of a cumulative MOE are discussed in a related study. A thesis examining the validity of the Sales Comprehension Test used this MOE with unsatisfactory results. The study reported that the cumulative MOE concealed a recruiter's monthly success rate. Since the MOE was a summary statistic, detailed analysis could not be performed in the study. As a result a second MOE, which reflects the basis of this study's monthly MOE, was developed. [Ref. 11, p. 26] A monthly approach presents a more accurate view of a recruiter's performance.

To ensure that the MOE for every recruiter was calculated from the same baseline, only the first nine months of a recruiter's recruiting time was used for the calculation. The first nine months was chosen for a number of reasons. Historically, an average of seventy-one percent of recruiters who fail because of ineffectiveness fail in the first nine months. [Ref. 12, p. 1] The first nine months is a new recruiter's Transitional Training and Evaluation (TTE) period. It is a period when the attributes he has **before** he became a recruiter would have the most effect. In a University of Illinois study by Humphrey on the Scholastic Aptitude Test (SAT), the correlation with a college student's success was highest in the first year of school, nine months in a regular school year. [Ref. 7, p. 346]

Adequacy of sample size was a primary concern. The sample size had to be large to ensure that the group under study was representative of all recruiters, allowing the results of the study to be applied to the recruiter population. A large sample set would also help establish the data's normality, which is essential to regression theory.

Recruiters are divided into three time groups. The first group consists of the TTE recruiters who have nine months or less of recruiting time. The second group are detail recruiters who have from ten to twenty-four months recruiting.

The last group are mostly career recruiters who have more than twenty-four months of recruiting time. A preliminary check showed that statistically, the mean success rate did not differ from group to group (Appendix C). However, the recruiters who qualified to have the minimum amount of mission months for each category necessarily varied in size. The number of recruiters who had a minimum of twenty-four months was less than forty. Career recruiters numbered even less. There were 101 recruiters who had at least nine months of recruiting. This group of 101 recruiters comprised the first workable database that could be used for the regression model (Appendix D).

Additionally, a frequency histogram of the monthly MOE, shown in Appendix E, gives strong indication of a normal distribution. A normal distribution suggests two points of discussion. Since the MOE functions as the dependent variable, regression assumptions require that its variability result only from residuals (error). If the dependent variable is normally distributed and its variability lies only in the residuals, then the residuals must also be normally distributed, which is a regression assumption that allows numerous statistical operations in regression, such as t-tests and F-tests. A less important point is that the MOE shows variability. A prediction of a constant makes the problem moot. A constant response variable suggests that no factors have an effect on the MOE. The monthly MOE gives evidence of being a logically and statistically good measure of recruiter success.

2. Explanatory Variables

a. Salesmanship Ability

Sales ability is a key factor for recruiter success. USAREC Manual 100-5 lists it as the number one factor for success. AR 601-1 says "...failure to develop the necessary sales ability to be a successful recruiter..." as grounds to identify a recruiter as ineffective. [Ref. 13, p. 14] Much of a recruiter's job involves aspects of salesmanship. He must know his product (U.S. Army) very well, and he must convince other people that it is a product that will benefit them. He must know his "market area", the environment where he must make his sales, and know how to relate to his clientele.

The SCT is a thirty-one question multiple choice test. The SCT was chosen because other tests were either too hard to administer, too complicated for the test subjects, too expensive, or applied to only specific groups of people. The SCT measures the degree to which an individual, from any category, comprehends basic sales principles. Scores on the SCT can range from a negative seventy-eight to a positive ninety-eight points. Questions do not have an absolute right or wrong answer, but degrees of right and wrong. This format makes it difficult to outguess the test. The test is strictly an aid to appraising success in a sales career. Its frequent use in many sales companies attests to its value as a predictor of sales success. [Ref. 2, pp. 2-3]

Construction of the test was preceded by extensive research and cross-validation in over 1,400 hundred cases. Normative charts which accompany the test show statistics for different groups. Distribution of SCT scores for sales students, sales clerks, and non-sales subgroups are presented. Statistical information for each group are also shown to emphasize differences or similarities with specific or general populations. Each group's mean has been shown to be significantly different from the means of other test groups, showing that different groups respond differently to the test. For instance, the mean score of a telephone sales clerk was 14, while the mean score of a hardware sales clerk was 29, and the mean score for non-sales personnel was 19. A key finding in test validation was the correlation between the final grades and test scores of students studying salesmanship in Rutgers University, Notre Dame, and St. Mary's College. The highest correlation coefficient was .70, which shows evidence that the test measured sales comprehension similar to that gained by students in sales school. [Ref. 2, p. 7] Scores on the SCT will be used as a measure of recruiter salesmanship ability.

b. Length of Service

A recruiter's length of service exposes him to some number of events, providing him with a certain amount of experience in the military. This experience may be advantageous to the recruiter when speaking with a prospect. His experience would allow him to present information about the military that the enlistee would find interesting. His previous dealings with young soldiers should give him an idea about what benefits, or attributes of the military might appeal to soldiers. Length of service would also represent the maturity subgroup that the recruiter should have. His maturity should make him more adaptable to new environments and enable him to cope with less-than-desirable situations.

A soldier's length of service may also indicate the amount of organizational skills he or she possesses. The longer a soldier stays in the military, the more chances he or she may have to develop organizational skills. Soldiers, at one time or another, assume a leadership position, more so for recruiter candidates who are in the grades of E5 and E6. A leadership role demands that the soldier exercise organizational skills to ensure that his unit accomplish tasks in a timely and efficient manner. Length of service will be measured in months, to coincide with the MOE. Information provided for the study included the months a recruiter had been in the service and the months he had been a recruiter. Since the study is interested in the length of service that the recruiter had before he became a recruiter, his months as a recruiter were subtracted from the months he had been in service when the database was constructed.

c. Armed Forces Qualification Test (AFQT)

Recruiters must possess an ability to communicate effectively with potential enlistees. Ninety percent of a recruiter's work involves exchanging ideas with enlistees and teaching them the benefits of the Army. To gain a commitment, the recruiter must be able to persuade. He must be able to present his product attractively. Versatility in verbal communication is essential.

The AFQT is a compilation of scores from the Armed Services Vocational Aptitude Battery (ASVAB), which was designed to measure the propensity of a soldier to perform well in a specific branch of the Army. The AFQT consists of the scores in Word Knowledge (WK), Paragraph Comprehension (PC), Arithmetic Reasoning (AR), and Mathematical Knowledge (MK). WK and PC are both verbal measures. The format of these tests concentrate on a soldier's ability to extract and analyze information from written sources. AR and MK are designed to measure the quantitative ability of a soldier. [Ref. 14, pp. 68-74]

Recruiters must be able to absorb and apply knowledge completely different from their PMOS in a relatively short period of time. Previous analysis has shown high correlation between AFQT scores and trainability. Soldiers who score well on the AFQT typically have higher scores in service schools. [Ref. 14, p. 30]

d. Non-Intellective General Factors (NIGF)

Some traits such as energy and enthusiasm are intangible, and may not be readily measured. Spearman's theory of NIGF offers ideas on how to use certain categorical traits of a recruiter candidate as indicators of enthusiasm and energy. These traits transform to determination, will, and drive for success in a military sense. A preliminary graphical analysis of each of the factors described in the following sections showed that there is possibly a significant difference between the mean success rates of each factor's subgroups (Appendix F). An expectancy chart was also constructed to determine the potential use of NGIFs as predictive variables (Appendix G).

(1) Primary Military Occupational Skill (PMOS). A soldier's PMOS may influence his attitude about duties as a recruiter. Recruiting command leaders feel that soldiers who have a PMOS from combat or combat support branches are more successful than recruiters who do not. Combat soldiers are felt to be more acclimated to pressures of mission-focused units. These soldiers are tested in mission accomplishment daily.

PMOS will be assumed to be an indicator of a recruiter's energy and enthusiasm. A combat arms recruiter would be expected to have the drive to complete the mission, regardless of the circumstances. A recruiter may not be content with recruiting duty, or may not possess a great amount of salesmanship, but still be determined to meet the mission. PMOS will be treated as a binary variable. A recruiter with a non-combat PMOS will be given a score of **zero**, while combat PMOS's will be scored with **one**.

(2) Selection. The process that assigns a soldier to recruiting duty may have an impact on performance as a recruiter. Soldiers are assigned to recruiting by volunteering or by being ordered by their military branch. The latter, and more common procedure, is called DA selection. Soldiers who are DA selected have been "volunteered" by their primary branch to fill a quota. DA selected soldiers have no options for a different assignment unless another soldier from the same branch wishes to swap assignments. USAREC has an unpopular reputation within the Army community. Scandals in the 1970's and early 1980's left many soldiers with the perception of USAREC service as career-ending duty[5].

[5]Recruiting scandals which resulted in the relief of many recruiters gave the Army's Noncommissioned Officer (NCO) community a very negative perception of recruiting duty and USAREC in general. The period after Vietnam added new pressures to meet enlistment quotas which resulted in some unscrupulous practices. The consequences of these questionable acts were routine, but were so wide spread that the long term effects still affect USAREC.

Volunteer recruiters want to be assigned to USAREC. Consequently, a volunteer recruiter may have more incentive than a DA-selected recruiter to perform well to continue his duty at USAREC. It is believed that a soldier who is content with his duty will be energetic and enthusiastic. Soldiers who have had no choice but to be a recruiter for three years could understandably be less enthusiastic and energetic in the performance of their duties.

The Recruit the Recruiter Program was conceived to eliminate negative perceptions of USAREC while increasing the number of volunteers. This program has found that volunteer recruiters have less of a propensity to fail in the Army Recruiting School (ARS) than do DA-selected recruiters. [Ref. 15, p. 1] The method of selection will be treated as a binary variable. A DA recruiter will be given a score of **zero** and a volunteer recruiter will be scored with **one**.

(3) Gender. Female recruiters may have an advantage over their male counterparts. Recruiting demands attention to detail and organizational skills. Every mission is accompanied by much documentation. The ability to manage administrative details is a necessity. Females may be more apt to coordinate these details. A secondary impact which may not be wholly intentional is the competitiveness that may exist for women in the military. A desire to prove themselves may drive them to work harder to meet mission. Gender will be treated as a binary variable. A female recruiter will be given a score of **zero**, and a male recruiter will be scored with **one**.

(4) Expectancy Charts. These NIGF's may or may not prove to be predictors for recruiter success. To support the use of categorical factors as predictive variables, an expectancy chart was constructed with success rates. According to McKenna, [Ref. 3, p. 61] if the categorical factor had a difference between its two subgroups that coincided with the expectancy

chart, then the factor may be considered a predictive variable. The calculated average success rate for all recruiters in the database was 0.6848. For each categorical factor's two subgroups the **proportion above** and below the **average** was computed. Distinguishing which subgroup is (1) and which is (0) is not relevant to the expectancy chart. Its main purpose is to determine if the categorical trait should be considered as predictive variables. The average success rates for each subgroup were calculated. If the average success rate for one subgroup was greater than the other subgroup, then the proportions above the average for each subgroup should also have the same pattern. If so, then the expectancy chart coincides and the categorical factor should be considered as a predictive variable. Table 1, a sample expectancy chart, is shown below for PMOS.

PMOS SUBGROUP	# ABOVE AVG	# BELOW AVG	% ABOVE SAMPLE AVG	SUBGROUP AVG SUCCESS
NON-COMBAT	33	24	0.5789	0.6987
COMBAT	20	24	0.4545	0.6669

Table 1. PMOS Expectancy Chart

The last column of the PMOS expectancy chart shows that the average success rate for recruiters with non-combat PMOS's are higher than recruiters with combat PMOS's. It should be expected that the proportion of recruiters with **noncombat** PMOS's who are above the sample average should be higher than the proportion of recruiters with **combat** PMOS's who are above the sample average. The fourth column of the chart shows that this expectation does occur, indicating that PMOS should be considered as a predictive variable. The expectancy charts for the other two NIGF's also support their consideration as predictive variables.

e. Factors Not Measured

Other measures of communication skills and organizational skills were considered as predictors of recruiter success, but could not be used. Measures of these traits were too subjective to include as measures in the analysis. One example is the Noncommissioned Officer Evaluation Report (NCOER). The NCOER requires a soldier's supervisor to give a rating on a scale of one to four on communication skills, as well as organizational and planning skills. Although these traits mark a successful recruiter, the context in which soldiers in their PMOS are rated is totally different from the environment of a recruiter. In a field unit, a soldier's communication skills is the ability to issue orders to subordinates. Communication skills in recruiting entail intelligent persuasion and conveyance of pertinent information. Planning in a field unit consists of being able to follow the daily training schedule. Recruiters must effectively plan their own schedule to optimize their productivity. These definitions are incompatible and so cannot be used for the study.

Other traits such as work habits and teamwork, which could be linked to recruiter success, were not measured. Limitations of time and resources prevented collection of measures for these and other factors. Not all traits of a successful recruiter could be linked to information in the database.

D. VARIABLE SYNOPSIS

This chapter has described the MOE which will be used as the dependent variable in the study and six variables that will be used as explanatory variables. The six independent variables initially chosen to predict recruiter success were as follow:

- AFQT = Score on the Armed Forces Qualification Test
- Test = Score on the Sales Comprehension Test
- Months = Soldiers length of service in the Army
- Gender = Male or Female

- PMOS = Combat or Noncombat
- Selection = Volunteer or DA assigned

Regression methods in Chapter III will show the separate and combined effects of these potential explanatory variables on the MOE.

III. VARIABLE SELECTION AND MODEL DEVELOPMENT

Because the extent to which a variable would have an effect on success was unknown, all the variables were treated as equally important. A simple linear regression for each of the six independent variables revealed the degree which it correlated with the MOE. Calculations of percent partial coefficients of determination showed the proportion of variability that each variable contributed to the full model. These calculations were preliminary steps in selecting the best subset of variables to use in the final model. Forward and backward regression analysis were initially considered for variable selection, but these methods were discarded because of undesirable consequences. Another method, based on the numerical value calculated from Mallow's coefficient, was used to select the best subset of variables for the initial model. Hypothesis tests, in conjunction with subsets from Mallow's coefficient, were instrumental in developing the final model.

A. CRITERION FOR "ADEQUACY" OF MODEL

To meet the predictive goals of the study, the formulated regression model must have at least a multiple correlation coefficient of 0.30. The correlation coefficient of the regression model was chosen as the primary criterion of model adequacy because it measures the degree of linear relationship between the combined effects of the predictive variables and the MOE. Knowing the value of one or more predictive variables allows an inference of a future value (the MOE) with some confidence, which is the primary purpose of the study.

This lower boundary for the criterion may seem low, but compared with the majority of screening instruments that try to predict future human behavior (in particular, success), it is adequate. Inevitably, the goal of the study is to predict a unique individual's future performance. Klitgaard cites from a Harvard study,

"...people are unpredictable..." [Ref. 17, p 4] He further reports that it is extremely difficult to predict the behavior of individuals because any performance criterion will have imperfect information. [Ref. 17, p 85] Ghiselli .es that in past studies of predictive tests, the correlations with the performance criteria have rarely exceeded 0.33, and more frequently have an average of 0.20. [Ref. 7, pp. 357-358] Combined verbal and math SAT scores, which are widely used to select students for entrance in colleges and universities, yield correlation coefficients averaging 0.41. [Ref. 16, p. 6] The verbal and math scores from the SAT have an average correlation with first year college GPA's of 0.38 and 0.34, respectively. [Ref. 7, p. 343] Robert Klitgaard, author of ***Choosing Elites***, also corroborates Ghiselli's results that the greatest correlation to be expected between predictive variables and actual job proficiency (rather than *training* performance) is about 0.35. [Ref. 7, p. 358]

It must be recalled that this study is an initial phase of a more general study of the recruiter selection process and is the foundation for continued research. Not all predictor variables will be included in this first model. As more factors which affect recruiter success are found, the multiple correlation coefficient will increase. Additionally, the database was not of optimum quality because of the likely homogeneity of the sample group with respect to the MOE. This homogeneity in the response variable tends to reduce the correlation with the explanatory variables. [Ref. 17, pp. 93-95] A correlation coefficient equal to or above 0.30 is a sufficient level for the final model in this initial study.

B. DATABASE INSPECTION

To ensure that the data set contained reasonable data, a scatterplot with the MOE was generated for each independent variable that was not a binary variable. Scatter plots for SCT scores and the AFQT indicated outliers (Appendix H). Outliers in one plot would not justify excluding the data. If the same recruiter caused outliers in different scatterplots, then the reason might be that in these particular instances, factors other than those in consideration are acting upon the dependent variable, and removal of the data points would be advisable.

Investigation of the scatterplots in Appendix H revealed two recruiters who were causing high leverage outliers in different plots. One recruiter had a very low score on the AFQT (17) and a negative score on the SCT (-1) while achieving nearly ninety percent success. The majority of recruiters who did poorly on the AFQT and/or the SCT had relatively poor-to-mediocre success percentages. Conversely, those recruiters who did well on the SCT and/or AFQT had average to superior success percentages. Another recruiter had the opposite situation. This recruiter did well in the SCT (24) and the AFQT (75) but was achieving less than forty-one percent success. Appendix H includes residual plots from simple regression calculations which show that the residuals of the success MOE for AFQT and SCT were greater for these recruiters than for most of the other recruiters. However, the plots also show that residuals from other recruiters were equally large or larger than the two recruiters in question.

Because these findings were inconclusive, and no additional information is available to determine if these recruiters are representative of the population or are unique data, they were kept in the database. Assessment of the success or failure of these recruiters as being due to factors other than those in the study cannot be readily explained. If more information on these recruiters indicated extraordinary conditions which would have affected their performance, the data would have been excluded. Until such information is obtained, the data will be included in the database and used for model development. However, a model was developed separately, using a database that excluded these recruiters, to determine the extent of difference that might result from the exclusion of these data (Appendix I).

C. SIMPLE REGRESSION OF EXPLANATORY VARIABLES

A simple regression presented a general idea of the degree of correlation between the success MOE and individual independent variables. Because of the small sample size, a ninety percent confidence interval was used. Each simple regression produced a Pearson's product-moment correlation coefficient:

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}}, \qquad \textbf{(Equation 1)}$$

where x_i is the i^{th} value of the independent variable, $\bar{x}$ is the average of the x values, y_i is the i^{th} observed dependent value, and $\bar{y}$ is the average of the y values. As a statistic, the variation of a correlation coefficient is a function of sample size. For this reason the correlation coefficient produced by simple regression in a sample does not portray a population r. However, as a descriptive statistic, it does show the relative strength of each independent variable's association with the MOE in the sample. This information can lead to the selection of an optimal subset of variables. Another statistic that allows insight to the importance of each variable to the model is the coefficient of determination, which measures the relative proportion of MOE variability accounted for by the independent variables. The coefficient of determination, r^2, is the square of the correlation coefficient and is also an output of the computer program's regression calculations. The resulting r and r^2 for each independent variable are shown in Table 2, below.

FACTOR	r	r^2
AFQT	0.1733	0.0300
Test	0.1474	0.0217
Gender	0.1641	0.0269
PMOS	0.1295	0.0168
Selection	0.1156	0.0134
Months	0.0233	0.0005

Table 2. Simple Regression Results

The r and r^2 values show that AFQT, Test, and Gender have good correlations with the MOE. Although PMOS, Selection, and Months may not at first glance appear to have a great deal of correlation with the MOE, it must be recalled that the study is working with real data and a small sample size. Real data almost never involve perfectly linear relationships, and a small sample size degrades the linearity that the data may show. These variables may still explain some of the MOE variability in multiple regression and should be considered for the final model. It can be reasonably expected that AFQT, Test, and Gender will be part of the final model. Months, PMOS, and Selection may still contribute significantly to it.

D. PARTIAL COEFFICIENTS OF DETERMINATION

The next step in determining appropriate independent variables for the final prediction model was to calculate the percent partial coefficients of determination, which were calculated by excluding one explanatory variable at a time from the full model and regressing the success MOE on the other five explanatory variables. The coefficient of determination of the reduced model was subtracted from the coefficient of determination of the full model and the result was in turn divided by the coefficient of determination of the full model. The equation,

$$\textit{Percent Partial Coefficient of Determination} = \frac{r^2_{full} - r^2_{reduced}}{r^2_{full}} X100$$

(Equation 2)

was used for each independent variable. This value represents the portion of variability that the variable can independently explain relative to the proportion of variability that the full model explains. The full model's r^2 had a value of 0.1186.

The calculation results are shown in Table 3.

EXCLUDED VARIABLE	REDUCED R^2	PERCENT PARTIAL R^2
AFQT	0.0956	19%
Test	0.1017	14%
Gender	0.0828	30%
PMOS	0.0907	24%
Selection	0.1025	14%
Months	0.1100	7%

Table 3. Percent Partial Coefficient of Determination

The listing shows that Gender, AFQT, and PMOS contribute the greatest proportion. These three variables, independent of their overlap, explain nearly seventy-three percent of the MOE variability that the ***full*** model can explain.

These results support inclusion of AFQT and Gender in the final model while Test's low percentage may suggest the adverse effects of multicollinearity. Additionally, it can be seen that PMOS contributes 24% to the overall fraction of variability explained by the full model. This fact may be evidence that PMOS is a factor that should be included in the final model. The relatively low percentages exhibited by Months and Selection are further indication that these factors may be excluded from the final model.

E. SELECTION OF BEST VARIABLE SUBSET

1. Forward and Backward Regression

The previous calculations have offered some valuable insight into the importance of each independent variable. There are a number of methods that can be used to select the variables to be used in the final model. Forward and Backward regression are the two most popular methods. Forward regression employs systematic addition of variables into the model. The procedure is simple. The initial model would be a simple regression that included the variable with the largest coefficient of determination. The model would be developed further by adding the variable to the initial model that produced a higher R^2 than adding any other variable would. This process continues until the difference in R^2, with and without a candidate variable, was below some previously determined value. Backward step-wise regression is similar, but begins with a full model and deletes one variable at a time.

These methods produce models with acceptable results, but they do not always select the subset of independent variables that optimizes the value of R or R^2. These methods fail to consider multicollinearity. Adding or deleting one variable may increase the importance of another, or vice versa. Additionally, there may be alternative models that produce the same value of R^2. Subjectivity may then play a part in variable selection, allowing one to virtually choose which model he or she *believes* is the most descriptive.

Multicollinearity of the selected variables is a special problem. Addition of a variable may cause coefficients of other variables to have less reliability. Deletion of a variable may cause bias so that the true value of the coefficient is systematically different from its estimated value. Recall, regarding multicollinearity, that the estimated variance for the j^{th} explanatory variable's coefficient estimator is calculated by the equation,

$$\hat{V}(b_j) = \frac{s^2}{(n-1)s_j^2} \times \frac{1}{(1-r_j^2)}, \qquad \textbf{(Equation 3)}$$

where s^2 is the error mean square (MSE), s^2_j the variance of variable j, and n is the sample size, while r^2_j is the coefficient of determination between the j^{th} explanatory variable and the other explanatory variables. From the equation, it can be seen that a high r^2_j results in a denominator with value close to zero, producing a large variance of the regression coefficient. Deletion of a correlated variable should decrease variability, but, depending on the data configuration, its deletion may introduce bias into the other regression coefficients. Additionally, it may cause the MSE to become inflated.

2. Use of Mallow's Coefficient, C_p

Use of Mallow's coefficient allows one to minimize the MSE of the estimate E(*y*) from the model. Using the error for E(y) as a criterion for model quality is a reasonable goal if the purpose of the model is prediction. [Ref. 18, pp. 18-19] Since the main goal of this study is prediction, Mallow's method suits its needs very well. The method considers the trade-off between decreasing error variance in the model by including or excluding variables and introducing bias into the estimated coefficients and multicollinearity among the variables. Mallow's method involves identifying all possible combinations of the independent variables being considered and calculating the coefficient C_p for each combination:

$$C_p = (m - p)(F_{m-p} - 1) + p, \qquad \textbf{(Equation 4)}$$

where m is the total number of independent variables being considered for the model, p is the total number of independent variables in the combination being evaluated plus the constant variable, and F_{m-p} is the statistic for testing the null hypothesis that the coefficients of the **excluded** variables are all equal to zero.

The F-statistic is critical in computing Mallow's coefficient. Recall that the F-statistic is calculated by dividing the sample variance of the regression model (the predicted Y's) by the variance (MSE) of the residuals. Therefore, when F_{m-p} is small, the MSE of the excluded variables is large. This is the desired outcome. If the excluded variables have a large MSE, then the variables included in the model are a good subset. The variables which would have caused a large MSE in the model have been eliminated, and the model will have a smaller MSE than with the presence of the excluded variables.

As the F-statistic decreases toward one (its median value), C_p approaches p. A forty-five degree line which is the line C_p = p indicates models with an F-statistic that is necessarily equal to one. These models have interchangeable excluded and included variable subsets.

However, a plot of C_p-versus-p in reference to the line $C_p = p$ is instrumental in selecting the best variable subsets which describe the MOE. Subsets above the line are not good subset selections because they have excluded variable sets with large F-statistics, meaning a small MSE in the excluded variables. Subsets below the line are good subset selections because they have excluded variable sets with small F-statistics, meaning a large MSE in the excluded variables. Concentration on these subsets below the line will lead to the initial model.

To obtain the C_p versus p plot requires calculation of the F-statistic for each possible excluded subset. Appendix J lists all the possible subsets, the variables which are included in each, their regression results, including the F-statistic for the null hypothesis that the p-1 coefficients in the regression equation were all equal to zero. The complement of the subset p is the set of excluded variables. Appendix K shows the F-statistic for the regression of the subset of **excluded** variables on the MOE. For instance, in computing C_p for the subset which included AFQT, Test, Gender, and PMOS (ATGP), the MOE was regressed on the excluded variables of Selection and Month. The resulting ANOVA for the regression is shown below in Table 4.

	df	SS	MS	F	p-value
Regression	2	0.0183	0.0092	0.7063	0.4959
Residual	98	1.2722	0.0130		
Total	100	1.2905			

Table 4. ANOVA for MOE Regressed on Selection and Month

It can be seen that the F-statistic is small (less than one), indicating a large MSE. This F-statistic, along with the values of m and p are used to calculate C_p. In this case the subset in consideration (ATGP) has four variables plus the variable constant. Therefore, p is equal to five. For all calculations m is always equal to six, the total number of potential predictive variables. The computation for this example follows:

$$C_p = (6 + 1 - 5)(0.7063 - 1) + 5 = 4.4126. \qquad \text{(Equation 5)}$$

Note that C_p is less than p which means that it will be below the C_p = p line. In this manner, the C_p for all potential subsets were calculated. Appendix K shows the list of C_p values, along with other information, used to produce Figure 1.

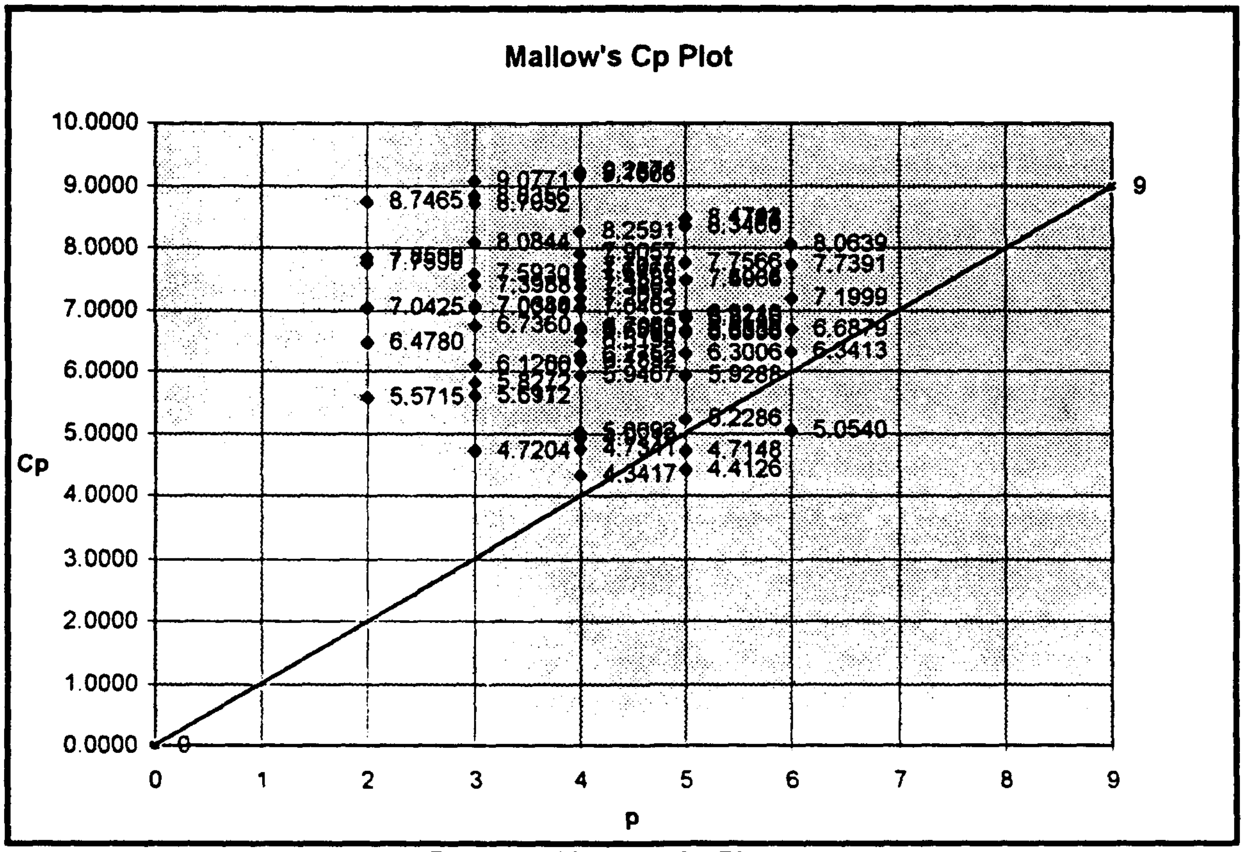


Figure 1. Mallow's Cp Plot

Each point in the plot is labeled with its Cp value which coincides with a numbered subset in Appendix K. The C_p plot shows three points which are below the line. They represent the subsets numbered 44, 56, and 60. These are the most promising subsets for the initial model. From Mallow's theory, these subsets will optimize R^2 and minimize the variance of the coefficient estimates and bias. [Ref. 18, p. 16] Table 5 below is a synopsis of the three subsets.

Subset Number	Factors	p	C_p	R	R^2
44	ATGP	5	4.4126	0.3082	0.0950
56	ATGS	5	4.7148	0.2866	0.0822
60	ATGPS	6	5.0540	0.3317	0.1100

Table 5. Mallow's Best Variable Subsets

From Table 5 and Figure 1 subset number 60, consisting of AFQT (A), Test (T), Gender (G), PMOS (P), and Selection (S) seems to best describes the MOE. As the subset with the largest multiple correlation coefficient, it will be used as the initial model.

3. Multicollinearity

High multicollinearity, or a large correlation coefficient between one explanatory variable and one or a combination of the other explanatory variables, may have undesirable effects. High multicollinearity makes the variances of the model's coefficients become large. It then becomes extremely difficult to distinguish the effects of one variable from another variable. Hypothesis testing on the coefficient estimators cannot be conducted with any high precision because the t-statistics tend to be very small and the confidence intervals very wide. These possible effects warrant a check of the final model.

A check for strong multicollinearity among variables is the variance inflation factor (VIF). The VIF is part of the equation of the estimated variance for the j th estimator, discussed previously. The latter part of the equation is the VIF and is computed by the equation,

$$VIF = \frac{1}{1-r_j^2}, \qquad \textbf{(Equation 6)}$$

where r_j^2 is the coefficient of determination of the j th explanatory variable regressed on the other explanatory variables. A VIF with a value of five or less is not considered to be significant. This figure corresponds to a r_j^2 value of 0.80. [Ref. 19, p.43]

Appendix L lists the VIF for each explanatory variable. All the VIF's are below 1.2, indicating that multicollinearity is not a concern. As discussed previously, inclusion of a variable into a regression model increases the correlation coefficient and exclusion of a variable decreases the correlation coefficient. Since the variable Month was **deleted** from the full model, the VIF's of the remaining variables can reasonably be assumed to be less than the previous calculations (<1.20), because the correlation coefficients of each remaining variable with the other four variables are smaller. Since this check does not warrant concern for multicollinearity, remedial steps were deemed unnecessary.

F. RESULTING MODEL

The procedures just described were used to select the initial subset of variables for predicting the MOE. A regression of the MOE on the selected subset would yield an equation of the form,

$$\boldsymbol{MOE = \beta_0 + \beta_A\,AFQT + \beta_T\,Test + \beta_G\,Gender + \beta_P\,PMOS + \beta_S\,Selection.}$$

where, **(Equation 7)**

- β_0 = 0.5368
- β_A = 0.0006434

- β_T = 0.0008856
- β_G =0.1070
- β_P = -0.03879
- β_S = 0.04889.

This initial model has a correlation with the observed MOE of 0.3317, which is adequate for the goals of this study.

G. DEVELOPMENT OF THE FINAL MODEL

Refinement of the initial model involves balancing the inclusion of variables having significant coefficient estimators with producing a model that has a multiple correlation coefficient within the goals of the study. Mallow's theory aids in obtaining this balance. Hypothesis testing can indicate the extent that each variable contributes to the model's predictive ability.

The initial model contains five independent variables: AFQT (A), SCT (T), Gender (G), PMOS (P), and Selection (S). A hypothesis test to evaluate if the coefficients of all the variables are truly zero is equivalent to the test that the coefficient of determination of the model is equal to zero. The null and alternate hypotheses are as follows:

$H_0 : \beta_A = \beta_T = \beta_G = \beta_P = \beta_S = 0$ **(Equation 8)**

H_a : At least one β_i does not equal zero; for i = A, T, G, P, S. **(Equation 9)**

If the null hypothesis cannot be rejected, then the model cannot explain any variations in the dependent variable and it is useless for the study. The significance level used is 0.10, an acceptable probability of a Type I error. The F-statistic, calculated from the ratio of the mean square regression (MSR) and the MSE, will be tested with the critical F value at the 0.10 level. The F-statistic must be greater than the critical F-value to reject the null hypothesis. The corresponding p-value should be no larger than 0.10.

The ANOVA table (Table 6) is shown below.

	df	SS	MS	F	p-value
Regression	5	0.1420	0.0284	2.3493	0.0467
Residual	95	1.1485	0.0121		
Total	100	1.2905			

Table 6. ANOVA Table for Initial Model

Since the p-value was 0.0467 (< 0.10), the F-statistic was large enough to reject the null hypothesis, which shows that the five variables in the model do explain at least some of the variation in the dependent variable beyond chance.

The next hypothesis test evaluates the significance of each of the estimated slope coefficients in the model. A t-test can be constructed to indicate if the variable, corresponding to the coefficient, contributes significantly to the model. The null hypothesis is that the coefficient is equal to zero; the alternate hypothesis is that the coefficient is not equal to zero.

$\mathbf{H_0}$: $\beta_i = 0$; for i = A, T, G, P, S **(Equation 10)**

$\mathbf{H_a}$: β_i does not equal zero. **(Equation 11)**

A t-statistic is calculated from the estimated coefficient, the null hypothesis value of the estimator (zero in this case), and the standard deviation of the estimator. This calculated t-statistic is then compared to a critical t-value at the 0.10 level with degrees of freedom (95 in this case) equal to the sample size minus one, minus the number of variables. To reject the null hypothesis, criteria similar to the F-test must be met.

Table 7, containing the t-statistic of each variable and a corresponding two-tailed p-value, is shown below.

	Coefficients	Standard Error	t Stat	p-value	Lower 90%	Upper 90%
Intercept	0.5368	0.0612	8.7760	0.0000	0.4352	0.6383
AFQT	0.0006434	0.0005	1.3513	0.1798	-0.0001	0.0014
TEST	0.0008856	0.0006	1.3948	0.1663	-0.0002	0.0019
GENDER	0.1070	0.0574	1.8629	0.0656	0.0116	0.2024
PMOS	-0.03879	0.0225	-1.7252	0.0877	-0.0761	0.0014
SELECTION	0.04889	0.0386	1.2671	0.2082	-0.0152	0.1130

Table 7. Initial Model Variable Statistics

From the table's p-value column, the variable that appears least likely to be a contributor to the model is Selection, followed by AFQT and Test. However, it is more appropriate to evaluate the AFQT and Test variable under a one-tailed hypothesis test, because the coefficients for these two variables are *expected* to be positive while the coefficients of the other variables can either be negative, or positive. Based on a one-tailed test, AFQT and Test are significant at the 0.0899 and 0.08365 levels, respectively. The t-statistic for the Selection variable indicates that the probability of rejecting the null hypothesis erroneously is greater than 0.20. The variable Selection will be deleted from the initial model and a new model, consisting of AFQT, Test, Gender, and PMOS, will be calculated and the model can be reevaluated.

This modification coincides with one of the variable subsets that Mallows method indicated would be a good variable subset to describe the data. The previous section showed subset ATGP as the second best choice. The resulting model is

$$\boldsymbol{MOE = \beta_0 + \beta_A\, AFQT + \beta_T\, Test + \beta_G\, Gender + \beta_P\, PMOS,} \qquad \textbf{(Equation 12)}$$

where,

- β_0 = 0.5377
- β_A = 0.0006361
- β_T = 0.0008636
- β_G =0.1106
- β_P = -0.03692

VIF calculations suggest no problems with multicollinearity. The previous hypothesis tests will also be applied to the new model. The resulting ANOVA table, Table 8, is shown below.

	df	SS	MS	F	p-value
Regression	4	0.1226	0.0306	2.5193	0.0462
Residual	96	1.1679	0.0122		
Total	100	1.2905			

Table 8. ANOVA Table for Refined Model

From the table it can be seen that the p-value meets the 0.10 crit[illegible]on set previously. Likewise, the separate explanatory variables in the ATGP model can be tested to evaluate their significance, using the t-statistic. Table 9 lists each variable coefficient's statistics.

	Coefficients	Standard Error	t Stat	p-value	Lower 90%	Upper 90%
Intercept	0.5377	0.0613	8.7650	0.0000	0.4358	0.6396
AFQT	0.0006361	0.0005	1.3320	0.1860	-0.0002	0.0014
TEST	0.0008636	0.0006	1.3564	0.1782	-0.0002	0.0019
GENDER	0.1106	0.0575	1.902	0.0575	0.0151	0.2062
PMOS	-0.03692	0.0225	-1.19	0.1042	-0.0743	0.0005

Table 9. Variable Statistics for New Model

It can be seen from the table that only the variable Gender is significant in a two-tailed hypothesis test. The least significant variable is AFQT, followed by Test and PMOS. But once again, employing a one-tailed test on AFQT and Test reveals that they are indeed significant. In Mallow's method, any subset, deleting AFQT, Test, or PMOS, is **not** a choice to describe the data optimally. Elimination of any of these variables decreases the model's ability to describe the variability in the MOE. The AFQT variable's relationship with the other remaining variables, for example, offset the relative weakness of its relationship with the MOE. Consequently, all the variables will be retained in the model.

The final model includes the variables ATGP and will be referred to as the ATGP model throughout the remainder of the study. The r^2 value of ATGP is 0.0950, and yields a correlation coefficient of 0.3082, which indicates that it is an adequate model to be used as a predictor of recruiter success for the purposes of this study.

H. THE FORWARD REGRESSION METHOD

The suitability of using the method of Mallow's C_p is demonstrated by employing another method to determine which variables should be included in the model. Using a significance level of 0.15 to evaluate the inclusion of each variable, a forward regression model was developed. This level is a common value used for this method. A lesser value would make the results questionable. Additionally, a model with a 0.10 significance level (the significance level used in Mallow's method) was evaluated. Predetermining the significance level was equivalent to setting a required level of change in R^2 because the F-statistic used for the significance test for each iteration can be calculated using the reduced and full models' R^2 's (full model refers to the model that includes the entering variable).

The null hypothesis is that the population coefficient of the entering variable is equal to zero. A simpler computation of the appropriate F-statistic is dividing the difference in the error sum of squares (SSE) of the full and reduced models by the MSE of the full model. If the resulting F-statistic is larger than the critical F value at the chosen significance level, then the null hypothesis is rejected and the variable in question entered the model. These calculations and iterations are shown in Appendix M.

At the a significance level of 0.10 (equal to the significance in Mallow's method) only the variable AFQT entered the model, resulting in an R equal to 0.1733 and a MSE of 0.0126. This model is not comparable to the ATGP model. To further the demonstration, a 0.15 significance level was chosen, resulting in a model that consisted of the variables AGP. The variable Test was not included. Because of the overlap of Test with the other variables, it did not produce enough of a change in the R^2 to justify its inclusion.

The advantages of using Mallow's method are obvious when the goal of the model is prediction. The focus of model development is to increase the correlation between the MOE and the predictive variables without adverse effects. The exclusion of Test decreased the model's R from 0.3082 to 0.2787, while just slightly increasing the MSE from 0.0122 to 0.0123. The trade-off, which Mallow's method emphasizes, is not present. In both cases, although small in one case, the outcome is worse. The AGP model is inferior to the ATGP model that Mallow's method indicates to be a good description of the success MOE. The appropriateness and usefulness of the ATGP model is the subject of the following chapter.

IV. MODEL DIAGNOSTICS, VALIDATION, AND APPLICATION

With the development of the final model, more scrutinizing procedures can be used to ensure that it indeed describes the system under study. Residual analysis will reveal if the linear regression model is appropriate, and data splitting will be used to cross-validate the model's predictive ability. A model that meets these necessary criteria will be considered a valid model, and will be used to calculate possible savings for USAREC from a hypothetical candidate group of recruiters.

A. RESIDUAL ANALYSIS

The model to be evaluated is

***Predicted Success Rate* = 0.5377 + 0.0006361(*AFQT*) + 0.0008636(*Test*) + 0.1106(*Gender*) - 0.03692(*PMOS*).** **(Equation 13)**

A study of the model's residuals will determine the appropriateness of a linear model to describe the data. The ATGP model's predicted success rates can be compared with the actual success rates of each recruiter. The difference between the two values are the residuals. Residual analysis focuses on validating the assumptions used to justify a linear regression approach. The main assumptions for linear regression are:

- error terms are independent
- error terms are identically distributed, normal random variables, $\varepsilon \sim N(0, \sigma^2)$; note that variance is constant
- the population regression model is linear in nature.

1. Independence of Error Terms

Independence of error terms refers to the assumption that the error in one observation does not affect the error in subsequent or previous ones. The most common occurrence of this phenomenon is time-series data. When data are collected in a time sequence, serial correlation may occur and violate the independence rule. As discussed in Chapter I, data were not collected in time sequence and the data are assumed to be independent. Additionally, t-tests on the success rates of the different time groups of recruiters were conducted, and are reported in Appendix C. The results showed no significant differences.

2. Normal Distribution of Error Terms

Normality of the residuals can be evaluated in several ways. Appendix N illustrates the three graphical methods used in this study. The upper left graph shows a frequency histogram. The contour of the residuals follows a bell-shaped pattern associated with a normal curve and is centered about zero, the theoretical mean of the residuals. The upper right graph of cumulative residuals matches the smooth, theoretical cumulative distribution of a normal curve.

Although the cumulative distribution of the residuals is rough, they closely follow the theoretical curve. The Kolmogorov-Smirnoff (K-S) bounds in the graph define the 90% confidence interval. The residual curve is within the bounds the entire length of the graph, suggesting that the cumulative distribution of the residuals follows a normal distribution. The last graph is a quantile-quantile plot. This graph plots the empirical residual percentile versus the theoretical percentile expected from a normal random variable. The more linear the plot is, the more indication that the random variable (residual) is normally distributed. The plot is almost perfectly aligned with the y = x line, except at the extreme ends, which may indicate unusual data points. These three illustrations provide strong evidence of normal distribution.

Mathematical calculations also support the assumption that the residuals are normally distributed. The residuals can be standardized using the equation,

$$Z = \frac{\varepsilon_i}{S_e}, \qquad \textbf{(Equation 14)}$$

where ε_i is the residual for observation i and S_e is the standard deviation of the residuals which have a mean of zero. Under standard normal conditions, with mean equal to zero and a constant variance equal to one, 95% of all residual observations should be within two standard deviations from the mean. [Ref. 20, p. 4-10] Table 10 below shows descriptive statistics of the residuals, some of which were used to transform the observed residuals into standard normal values.

ACTUAL RESIDUAL SUMMARY STATISTICS	
MSE	0.01220
Mean	0.0000
Standard Error	0.0108
Median	0.0038
Standard Deviation	0.1104
Sample Variance	0.0117
Skewness	-0.0824
Range	0.4986
Minimum	-0.2721
Maximum	0.2265
Count	101.0000

Table 10. Actual Residual Statistics

The standardized residuals are calculated from the observed residuals using the previous equation and information from the table. The percentage of standardized residuals which were within two standard deviations of the mean was over 97%, exceeding the 95% criterion. Appendix O lists the standardized residuals. Almost every standardized residual is between the values $\pm$ 2. The table also highlights other facts. The observed residuals have a mean of zero, and the minimal skewness of the data indicates that the distribution is symmetrical. Both are characteristics of a standard normal distribution.

The last investigation used to test for normality is the Chi-Square test for goodness-of-fit. The Chi-Square statistic is the sum of squared normal random variables. The hypothesis tested is that the random variables used to calculate the Chi-Square statistic comes from a normal distribution. If the hypothesis is true, then the statistic computed will not be an unusual value for a Chi-Square distribution. A large p-value (> 0.10) would be desired. A very small p-value would indicate that the statistic is unlikely to be observed in a Chi-Square distribution that represents a normal variable. Therefore, a large p-value would force the acceptance of the null hypothesis because there would be a high probability of being wrong if the null hypothesis were rejected. [Ref. 21, p. 197]

Appendix P presents the final statistics of the Chi-Square test. The degrees of freedom used for the test are the number of intervals between plus and minus infinity, minus one, minus one for the estimated mean, and minus one for the estimated variance. The number of divisions between plus and minus infinity is calculated using the equation 1+ln(2n), where n is the number of data points.[6]

For 101 data points there will be seven divisions, resulting in eight intervals, which is corroborated by the Chi-Square Goodness-of-Fit Table in Appendix P. The degrees of freedom is 8 - 3 = 5.

The columns *Observed* and *Expected* in the Goodness of Fit Table contains the data from which the Chi-Square value is obtained. The *observed* number, which is the random variable, is the number of data points that actually fall within an interval, and the *expected* number are the data points which were expected to fall within an interval, if the variable is normally distributed. The resulting Chi-Square from the residual data is 3.9781 with a p-value of 0.5526, which shows that this value is not unusual for a Chi-Square distribution. It

[6]The Chi-Square test was conducted using a NPS netware called A Graphical Statistical System (AGSS). AGSS takes the natural logrithm of twice the number of data points, adds one, and rounds up to the nearest whole number. This explanation can be found in the HELP section within the program.

indicates that the residuals have a normal distribution, because there would be over a 55% chance of being wrong for rejecting the null hypothesis that the random variables came from a normal distribution. It can also be seen that the statistic is near the Chi-Square mean, which is its degrees of freedom, lending further evidence that the statistic is not unusual for a Chi-Square distribution.

3. Linearity

Linearity of the model can be determined using purely graphical means. A scatterplot of standardized residuals versus predicted success rates is a key graphic, as shown in Figure 2.

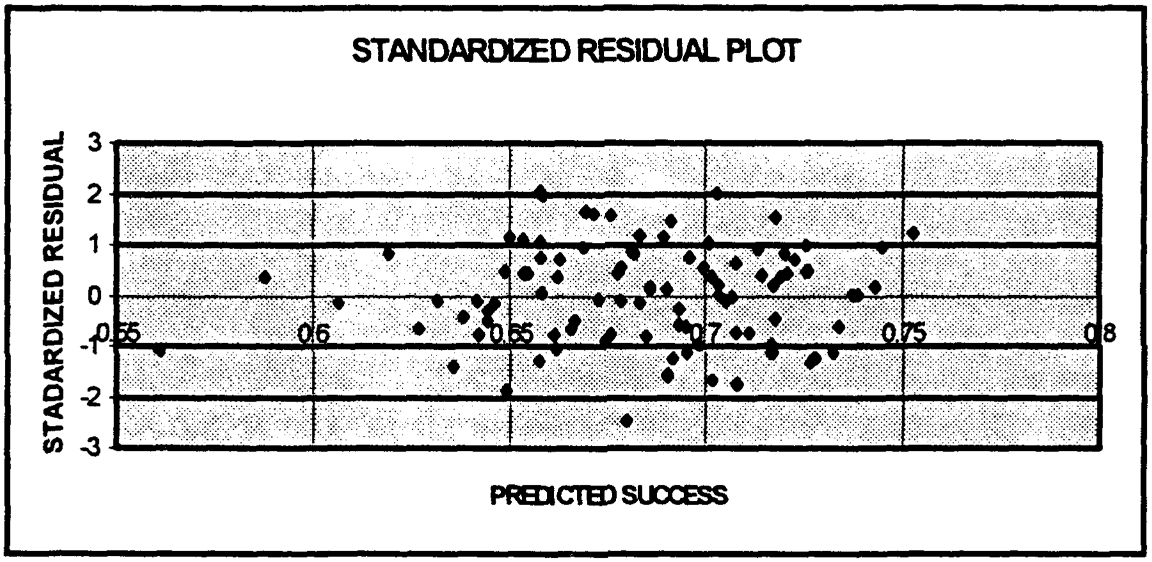


Figure 2. Standardized Residual Plot

Linearity is interpreted from the dispersion of the residuals. Any systematic pattern, such as an increase or decrease in the amount of error as the predicted value changes, means nonlinearity. As the plot shows, there is no pattern. The residuals remain equally dispersed about the mean zero throughout the range of the predicted success rates, indicating that the data follow a linear model. [Ref. 22, pp. 118-119]

The standardized residual plot also emphasizes other facets of the model. As previously stated, the errors remain in a constant pattern about the mean zero. This fact supports the assumption that the error variance is constant. [Ref. 22 p. 120] Additionally, outliers can be seen from the plot. From the discussion of normality, a standard normal plot should show that 95% of the observations are within 2σ. Residuals falling outside of the range should be investigated.

Although the plot does have data points outside the range, there is insufficient information in the database to suggest that these data points are not due to extraneous conditions.

Residual analysis has validated the assumptions justifying the use of linear regression. As a result, it can be said that the final ATGP model is an appropriate model for describing recruiter success.

B. MODEL VALIDATION

Validation of the model is the final step in model development. Validation examines the model's ability to perform within the bounds it was designed. The best way to test the model is to check it with new data. However, acquiring new data is costly, time-consuming, and requires resources not available for this study. Cross-validation is an option which uses present data to validate the model. It allows part of the present data to be treated as if it were new data. This method involves data splitting and employment of several evaluation techniques to ascertain validity.

1. Data Splitting

The entire data set is split into two separate sets: the model-building set and the validation set. It is important for the model-building set to be sufficiently large. The number of cases for the model-building set should be *at least* between six and ten times the number of variables in the model. [Ref. 22, p. 467] The number of cases in the model-building set will be 71, approximately eighteen times the number of variables. The number of cases in the validation set will be 30, the remainder of the database.[7]

The data were split by assigning each observation a number from 0 to 1 from a random number generator. The database was then sorted in ascending order; the first 71 cases were assigned to the model-building set with the remainder assigned to the validation set.

2. Cross-Validation

Two methods of cross-validation were employed to validate the model. The *mean squared predictive residual* (MSPR), computed from the validation set, can show the accuracy of the model in comparison to its expectations from the model-building set. The correlation coefficient between the predicted success rates, using parameter estimates from the model-building set's ATGP model and the actual success rates within the validation set can indicate if the model performs as well as expected.

a. Calibrating the Model's Predictive Ability

The model determined from the model-building set is

$$\textit{Predicted Success Rate} = 0.5422 + 0.000711(\textit{AFQT}) + 0.000566(\textit{Test}) + 0.1263(\textit{Gender}) - 0.0313(\textit{PMOS}). \qquad \textbf{(Equation 15)}$$

Using this formula, an estimated success rate ($\hat{Y}$) for each recruiter in the validation set was calculated. The difference between the estimated and actual success rates (Y_i) of each recruiter produced a residual. The MSPR was then computed from the equation,

$$MSPR = \frac{\sum_{i}^{n^*}(Y_i - \hat{Y}_i)^2}{n^*}, \qquad \textbf{(Equation 16)}$$

[7]Using a 70/30 percent split of the data for cross-validation was suggested by Professor Ronald Weitzman, System Management Department, as good analytic practice.

where n^* is the number of recruiters (30) in the validation set. The MSE of the regression model from the model-building set can be compared with the MSPR. If the MSE of the model-building set is fairly close to the MSPR, then the model shows good predictive ability. [Ref. 22, p. 466] Calculations for the estimated success rates, the resulting residuals from the validation set, and the MSPR are in Appendix Q. The MSPR was 0.01358, and the MSE of the regression model from the model-building set was 0.01219. The MSPR is within 11% of the model-building set's MSE, validating the model's predictive ability.

b. Correlation of Predicted and Actual Success Rates

Another technique to validate the model's predictive ability is to compute the correlation between the actual success rates within the validation set and the predicted success rates. The resulting correlation coefficient should be close to the correlation coefficient for the model-building set. The correlation is expected to be lower than the full model's correlation coefficient (n = 101) because of "shrinkage", but should not be unreasonable. Shrinkage reflects the inefficiency of a model when it is applied to new data. In this case the model-building set had an R of 0.3308. Figure 3 shows the scatterplot of the data.

A simple linear regression of the actual success rates on the predicted success rates can provide the statistics needed. The coefficient of determination has been included in the plot. The value of R^2 is 0.1530, which is equivalent to an R of 0.3911. The R of the final predictive model, being evaluated is 0.3082. Although shrinkage was expected, it did not occur. Instead, the correlation coefficient was greater than the correlation coefficient of

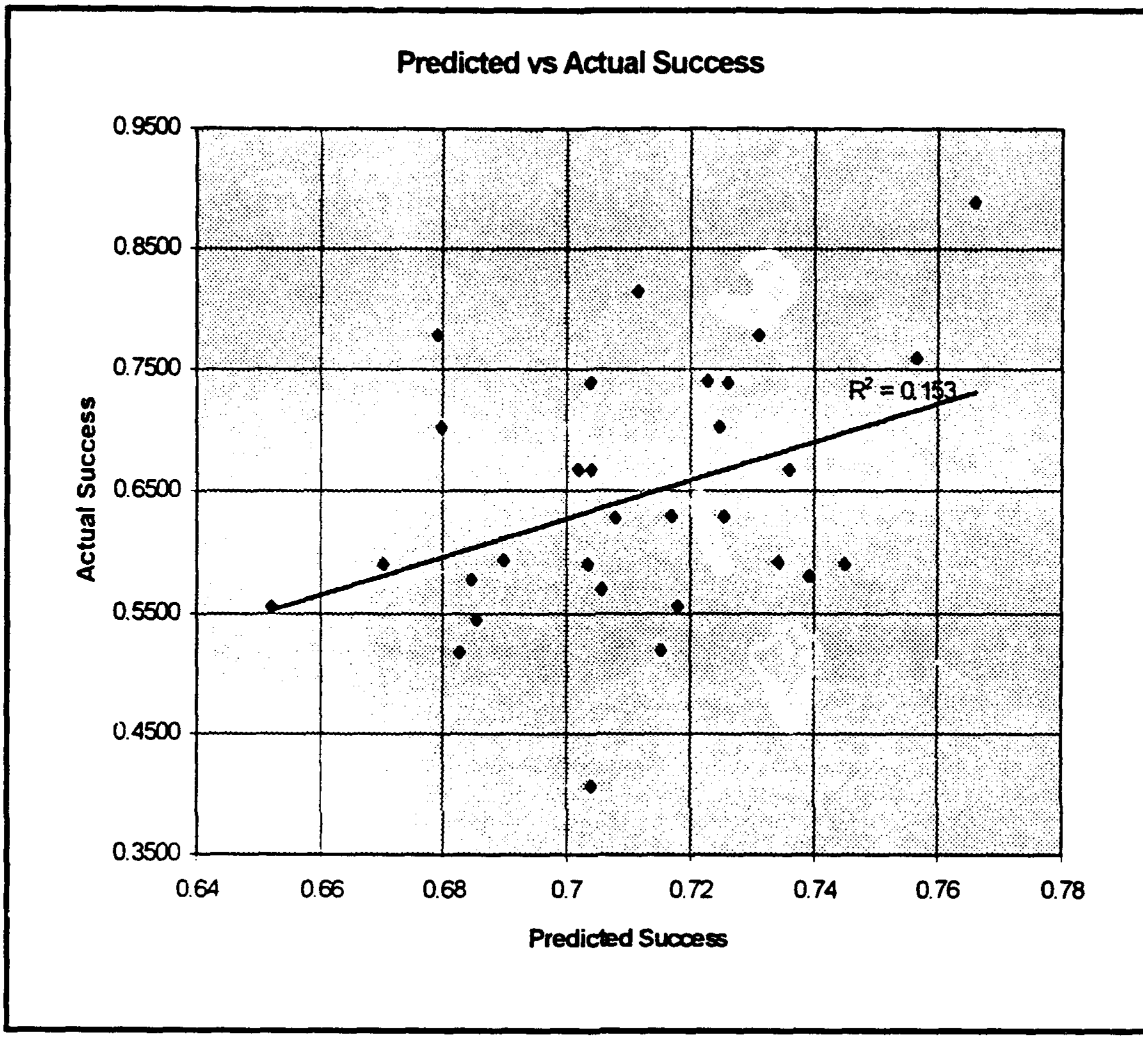


Figure 3. Actual Success vs Predicted Success

the ATGP model from either the full data set or the model-building set. This result demonstrates the model's ability to perform within its expectations.

These results from both methods of cross-validation support the model's ability to predict success. Validation of the model shows promise for the model's application.

3. Unification of the Data Set

The previous methods have validated the final model to be used. The correlation coefficient between the predicted and observed success rates from the validation set supported the model's ability to perform. The model-building set and the validation set will be merged and the original model will remain unchanged. The final formula which will be used to predict recruiter success is

$$\textit{Predicted Success Rate} = 0.5377 + 0.0006361(\textit{AFQT}) + 0.0008636(\textit{Test}) + 0.1106(\textit{Gender}) - 0.03692(\textit{PMOS}). \quad \textbf{(Equation 17)}$$

C. CALCULATION OF THE PREDICTION INTERVAL

A point estimate of a recruiter's success rate does not account for possible error in the estimate. A prediction interval (PI) is used to allow for error. A PI is computed using the square root of the regression model's MSE (*S*) and a predetermined significance level ($\alpha = 0.10$). A PI is different from a confidence interval (CI) in that a PI refers to possible values of individual Y measurements, rather than expected Y measurements. The interval accounts not only for the variance in the regression model (which a CI does), but also the variance of the error components (ε). This consideration results in a wider interval than a confidence interval, which only accounts for variance in the regression model (from sample to sample).

Since the model is a multiple regression model which reflects changes in more than one variable at a time, the PI depends on a vector of measurements. [Ref. 23, p. 532] The formula for the prediction interval is

$$PI = \hat{Y}_i \pm t_p S_e \sqrt{1 + a^T (X^T X)^{-1} a}\,, \qquad \textbf{(Equation 18)}$$

where ***a*** is the vector of independent-variable measurements used to compute a predicted MOE and $\boldsymbol{a}^T$ is its vector transpose. It should be noted that the ***a*** vector is a 5 x 1 vector, where the first row is a constant value 1 that applies to the constant in the model equation. The ***X*** in the equation is the matrix of measurements on the four independent variables used to develop the model, and $\boldsymbol{X}^T$ is its transpose. The ***X*** matrix is an n x 5, with the first column being a column of ones. Matrix calculations are in Appendix R. The t-statistic in the formula uses the number of points (individuals) in the database minus the number of coefficients, including the model's intercept, as the degrees of freedom (df), and p is the significance level used for the interval. The regression model has used a sample size of 101. Therefore the df is 96. For a two-tailed prediction interval, α would be divided by two, which gives p a value of 0.05.

However, the concern in selection is to avoid **overestimating** a candidate recruiter's success. A worse-case scenario implies that a one-tailed interval, with the cut tail on the lower side, is more suitable. The estimates thus obtained from the PI equation are now lower bounds. There is 90% confidence that the actual success rate of a recruiter candidate will not fall below such a bound. A recruiter who has an actual success **greater** than his or her expected y measurement does not negatively affect the system. Therefore, a higher bound is unnecessary, and the value of p will be 0.10.

At the 0.10 significance level, the value of $-t_{0.10,\,96}$ is -1.292, and the square root of the MSE is 0.1103. The resulting PI is

$$PI = \hat{Y}_i - 0.1425\sqrt{1 + a^T (X^T X)^{-1} a}\,, \qquad \textbf{(Equation 19)}$$

where $(X^T X)^{-1} =$

$$\begin{vmatrix} 0.309367 & -0.00081 & -0.00069 & -0.24595 & -0.0007 \\ -0.00081 & 1.87E-05 & -4.7E-06 & -0.00021 & 2.44E-0 \\ -0.00069 & -4.7E-06 & 3.33E-05 & 0.000231 & -1.8E-0 \\ -0.24595 & -0.00021 & 0.000231 & 0.272188 & -0.019 \\ -0.00073 & 2.44E-05 & -1.8E-05 & -0.0192 & 0.04163 \end{vmatrix}$$

As previously discussed, the SCT was administered to a BNCOC class, and pertinent characteristics of the students were collected to establish a test base which consisted of 46 complete records. Using this information and the recruiter success model, a predicted success rate was calculated for each BNCOC student. From the PI equation for each BNCOC student, the lower bound of each student's predicted success rate interval (PSRI) was determined. Appendix S presents these lower bounds. While the last student has a predicted success rate of 0.567, for example, his lower bound is 0.396.

D. POTENTIAL USAREC SAVINGS

This section will explore possible cost savings that use of the recruiter success model may generate for USAREC.

1. USAREC Losses

A current study by USAREC for FY90 showed that in one year 403 TTE recruiters were released from the command. Of these 403 recruiters, 382 were relieved for ineffectiveness and/or for being unqualified. The total cost to USAREC for these 382 recruiters was $4,693,828.80. [Ref. 24, pp. 1-3] The average cost breakdown for each recruiter is shown in Table 11.

ITEM	COST
Travel/Per Diem for ARS	$2250.00
ARS Training: 4 Weeks	$1425.00
SDAP[8]: First 3 Months	$495.00
SDAP: Last 6 Months	$1320.00
Recruiter Expense Allowance	$354.51
Vehicle Cost	$2983.00
Equipment and Supplies	$948.00
Clothing Allowance	$362.00
Station Commander Training	$2150.00
Total Cost for One Recruiter	**$12,287.51**

Table 11. Cost to Maintain Recruiter: ARS-TTE

2. Model Application

The usefulness of the final predictive model lies in its ability to identify undesirable recruiters before the expenditure of funds.

There are a total of 39 recruiting classes held every year at the U.S. Army Recruiting School, each class having an average of 40 students[9]. The average total number of new recruiters a year is approximately 1560. If the 382 recruiters

[8]According to USAREC PAE. SDAP stands for the Special Duty A ent Pay that all recruiters receive for being assigned to duty not within the scope of their PMOS.

[9]Major Alan Poikenen, Chief Analyst for USAREC PAE, provided these figures as an estimated average for the ARS. They are estimates because of the continually changing force structure in the U.S. Army.

who failed during the TTE period are the typical number of TTE failures each year, then an expected fraction of failures for each class can be computed.

Calculating this fraction requires dividing the total number of failed recruiters in one year (382) by the yearly average number of new recruiters entering the command (1560). The calculation reveals that the expected percent of TTE failures from one year's group of new recruiters is 24.5%.

The Taylor and Russell Tables, discussed in Chapter V, are of great use in cost-benefit analysis. Appendix T shows these tables. To properly employ these tables, it is necessary to define the terminology associated with them.

- **Base Rate (BR)**: The proportion of a population that is determined to be successful by the "employer".
- **Selection Ratio (SR)**: The proportion of applicants that are chosen from the population.
- **Correlation Coefficient (R)**: The validity coefficient of the predictive tool.(For this study's purpose, it is the multiple correlation coefficient of the multiple linear regression model R.)
- **Proportion of Expected Success**: The proportion of the selected applicants that is expected to be successful.

The following steps apply the Taylor and Russell tables to a cost-benefit analysis of the PRiSM in the current USAREC environment.

<u>**STEP 1**</u>: Determine the population from which the candidates come. The population in question for USAREC are NCO's in the grade of E-5 or E-6. Since NCO's of these grades must attend PLDC and/or BNCOC, it is possible to obtain an approximate yearly number of new NCO's in these grades. The Defense Manpower Data Center estimates that annually, 12,285 students attend BNCOC

and 10,000 students attend PLDC. Thus, the population that USAREC draws their recruiters from has approximately **22,285** members.

STEP 2: Determine BR. It has been determined that 24.5% of all new recruiters fail during the TTE period. Therefore, 75.5% of new recruiters are successful during the TTE period. The BR is the proportion of the population that is successful. If this group of new recruiters is representative of the population, then it can be said that **75.5%** of the population is successful, and 0.755 will be considered the BR.

STEP 3: Determine SR. USAREC admits approximately 1560 new recruiters into the command each year. USAREC draws from a population of 22,285 new NCO's each year. SR is computed by dividing 1560 by 22,285. The resulting SR is **0.07**.

STEP 4: Determine the Correlation Coefficient. The PRiSM has an R of **0.3082**.

STEP 5: Determine the Proportion of Expected Success. Each BR has a separate table. Since our BR is 0.755, enter the table closest to this fraction. The closest BR table is 0.80. The table is arranged with R on the left-most column and the SR on the top row. The intersection of R and SR is the Proportion of Expected Success. Part of the table for a BR of 0.80 (Table 12) is shown below.

r	Selection Ratio					
	[redacted]	[redacted]	0.20	0.30	0.40	0.50
0.25	0.92	0.91	0.89	0.88	0.87	0.86
[redacted]	[redacted]	[redacted]	0.90	0.89	0.88	0.87
0.35	0.95	0.94	0.92	0.90	0.89	0.89
0.40	0.96	0.95	0.93	0.92	0.90	0.89

Table 12. Portion of Base Rate Table

Entering the table with an R of 0.3082 and a SR of 0.07, results in an interpolated Proportion of Expected Success of **0.932**. Of those selected, **93.2%** are expected to be successful. With these results, cost-benefit analysis can be conducted. Comparing the numbers who would fail with and without PRiSM as a predictive tool makes it possible to compute potential savings. If the comparison is not significant, then the benefits of the tool may not be worth the costs.

Currently, 75.5% of all recruiters are successful, during the TTE period. If the PRiSM were used, then 93.2% would be expected to be successful. The corresponding failure rates are 0.245 and 0.068. Instead of 382 failures (0.245 x 1560) annually, use of the PRiSM would result in only 107 failures. Since 382 - 107 = 275, the annual savings from the use of PRiSM would be

$$275 \times \$12{,}287.51 = \$3{,}379{,}065.30. \quad \textbf{(Equation 24)}$$

3. Cost for Model Application

There are costs associated with using this model, but they are a minute fraction, compared to the savings. The Cost Effectiveness Ratio (CER) gives an idea of the monetary value of the prediction model. The CER is computed by dividing the cost of applying the model with the cost savings it produces. The ASVAB is a test already administered to all soldiers entering the military. The AFQT is formed from sub-scores of the ASVAB. Except for SCT scores, all other information used in the recruiter success model require no additional effort or funds.

Model application requires SCT score data, which is the only new cost to the recruiter selection process. The Sales Comprehension Test costs $200 for every 60 tests, an average cost of $3.33 per test[10]. Annually, USAREC admits approximately 1560 new recruiters. If the SCT were administered just prior to entering ARS, then all 1560 new recruiters could be screened at an additional cost to USAREC of $5194.80.

From an overall perspective, the cost of administering the SCT depends on when it will be administered. For example, if every soldier is tested at service entry, then there will be many, many more individuals tested, than if only potential recruiters were tested. Consequently, the time when it will be given to potential recruiter candidates changes the number of soldiers who will take it, and will in turn affect the total cost of the test.

For the 22,285 PLDC and BNCOC members, the cost would be 22,285 x $3.33, or $74,209.05 -- a small fraction of the savings to be expected from the use of PRiSM.

[10]Information available to the Naval Postgraduate School's Operations Analysis budget department shows that tests must be bought in bundles of 60 at a cost of $200 per bundle.

V. CONCLUSIONS AND RECOMMENDATIONS

This thesis sought to determine the feasibility of developing a predictive model that could improve the United States Army Recruiting Command's recruiter selection process. Intermediate goals were defining recruiter success in clear, quantitative terms and identifying traits that may indicate an individual's potential for success in the recruiting arena. A valid model may be instrumental in reducing the number of nonproductive recruiters in USAREC and improving the overall performance of the command.

A. MODEL LIMITATIONS

A shortcoming of the PRiSM, of which a user should be aware when considering its employment, is the quality of the database used to develop it. As discussed in Chapter I, the data that was available for the study was not optimal. Because data was collected using the present-employee method, the database lacked information for recruiters who were not successful. As a result, the recruiters used for the study were on the "right tail" of the recruiter population. These recruiters were all considered successful. This fact suggests that discretion should be used when interpreting the degree that any one predictive variable effects the success rate. It is possible that numerous recruiters who were unsuccessful could have had high scores on the AFQT and the SCT. The possibility of negative coefficients for these variables could then be raised.

The lack of variability in the recruiters may also account for the relatively low correlation between the individual predictive variables and the MOE. Although measurements of the independent variables had good diversity, the success rates of the recruiters remained at the upper end of the success scale. Scores for AFQT ranged between 17 and 99, while SCT scores varied from -23 to 58. Conversely, 91% of the recruiters had success rates above 0.55, while over 71% had success rates above 0.60, indicating somewhat low variability in the MOE. Until data that is more representative of the recruiter population is captured, the PRiSM should be used with prudence. Included in these discussions are suggestions for optimally employing the PRiSM within its limitations and recommendations to improve the model through more effective data collection and experimental design.

B. CONCLUSIONS

Through regression methods, a predictive model was developed that incorporates traits of a successful recruiter. The primary use of this model is in identifying potentially nonproductive recruiters who would eventually be relieved from recruiting at an enormous cost in funds, lost prospects, and unit efficiency. Screening these candidate recruiters before any USAREC funds are expended is the key advantage in having such a model.

The Predictive Recruiter Success Model (**PRiSM**) identifies four traits as predictors of an individual's recruiting success, which is measured as the fraction of TTE months that he will make mission. The final equation is

$$\textbf{\textit{Predicted Success Rate}} = \beta_0 + \beta_A X_A + \beta_T X_T + \beta_G X_G + \beta_P X_P$$

where, **(Equation 25)**

$\beta_0 = 0.5377$

$\beta_A = 0.0006361$

$\beta_T = 0.0008636$

$\beta_G = 0.1106$

$\beta_P = -0.03692$

X_A = Armed Forces Qualification Test (AFQT) Score

X_T = Sales Comprehension Test (SCT) Score

X_G = Gender (Male = 1, Female = 0)

X_P = Primary Military Occupational Skill (PMOS) (Combat = 1, Noncombat = 0).

The correlation coefficient for the final model is 0.3082.

Both quantitative and qualitative variables are part of the recruiter success model. Scores in the AFQT and the SCT, respectively, measure a soldier's communicative and potential sales skills. Gender and PMOS are categorical variables differentiating groups with perhaps greater or less potential for success in recruiting.

The numerical sign of each coefficient shows that those who have higher scores on AFQT and SCT tend to have greater recruiter success. On the average, male recruiters are more successful than their female counterparts. However, because of the small sample size, only four females were in the database. Female recruiters comprise ten percent of the active recruiters in USAREC, but the sample had only four percent. Although statistical tests yielded evidence that female recruiters were not as successful as male recruiters, a larger number of female recruiters would be needed to corroborate this finding.

A surprising result was the negative effect of having a combat PMOS. Many recruiting leaders believe that soldiers with combat PMOS's are more likely to be successful. This analysis in the study shows evidence to the contrary. Soldiers with noncombat PMOS's are more likely to be successful in recruiting.

Of the four predictive variables, three are insignificant at the 0.10 level when using a two-tailed hypothesis test. The two variables of most concern are AFQT and SCT. The two-tailed p-value for AFQT is 0.1860 and 0.1782 for SCT. However, it is expected that higher scores on both these tests should correspond with greater recruiter success. It would be unreasonable to expect that good communication skills and sales ability would hinder a recruiter from being successful. Hence, a one-tailed test, indicating that the coefficient should be positive, is more appropriate. For this test, the p-values are 0.093 and 0.0891 for AFQT and SCT , respectively, showing that the variables AFQT and

SCT are indeed significant. Lastly, the PMOS variable has a p-value of 0.1042, when using a two-tailed test, which indicates that the PMOS variable may have some significance. Although it can be argued on the basis of p-values that PMOS should be excluded from the model, this paper asserts that there is evidence to support PMOS as a meaningful predictive variable.

Because the information used in the study was collected from a survey, and not through experimental design, it would be extremely difficult and most likely faulty to infer causality from the study's results. Interpreting a recruiter's AFQT and SCT scores to be the direct cause of his success would be a mistake, because the recruiters were selected independently of these scores. In light of this argument, it may be possible to glean some insight on the inclusion and exclusion of some of the variables which were considered for the model.

A variable for a soldier's months in service was tried, but not included in the model. It had been thought that experience in the military would be an advantage for a recruiter. Because the recruiting environment is so different from most soldiers' primary duties, it is possible that whatever experience the soldier does possess is nullified when entering recruiting duty. When a soldier starts his recruiting assignment, he begins at the same baseline as all the other recruiters. His ability to absorb the training he receives in recruiting school and perhaps natural ability in recruiting may enhance his chances for success.

With these facts in mind, it may be reasonable to understand why AFQT and SCT scores were variables that were included in the model. AFQT scores not only indicate a soldier's communicative ability; they also are a good measure of intelligence. Thus, an intelligent soldier should comprehend the recruiter training he receives and will more aptly apply this knowledge than a soldier who has not understood recruiting instructions so well. Since most of the duties of a recruiter involve many aspects of salesmanship, an understanding of basic sales skills is an advantage for a new recruiter. It is this same understanding that the SCT measures and recruiting duty demands.

Another variable that was tried, but ultimately not included in the model, was Selection. The method that a soldier was assigned to recruiting duty was thought to have an effect on his or her performance. The Recruit the Recruiter Program (RRP) was formed to increase the number of volunteer recruiters in USAREC, a goal that it is successfully achieving. In accomplishing this task, the RRP has not used the key traits common in successful recruiters to mprove its selection criteria. It continues to use the administrative guidelines in AR 601-1, which do not make any reference to the characteristics associated with success outlined in USAREC Manual 100-5. Accordingly, on the average, recruiters who have volunteered for recruiting duty fare no better than recruiters who have been nominated by their primary branch. An example of this indifference to the method of selection can be found in the Air Force Recruiting Command.

Although the Air Force has 100% volunteer recruiters, it still has the same problems that USAREC has been experiencing[11]. The exclusion of the Selection variable supports this finding.

A supplementary finding was the effect of adding a Unit variable to the final model. Although this study found no significant differences between unit success means during the nine month TTE period, it was felt that this variable would increase the predictive ability of the model. The addition of a Unit variable increased the correlation coefficient from 0.3082 to 0.3300. These calculations and the modified model are in Appendix U.

Including this variable had positive effects, but it is impractical. It is improbable that a candidate recruiter will know where he will be assigned. By the time an assignment is designated for the recruiter, the PRiSM will have been

[11]During the Joint Manpower and Recruiting Conference held in Rockville, Maryland on 29 June, 1994, Mr. George Germadnik from the Air Force Recruiting Command discussed the similar problems that the Army and the Air Force are experiencing with recruiter attrition.

used, the candidate will have been selected, and funds will have been spent. If the PRiSM were to indicate negative results, these expenditures will have been wasted. Since there are forty recruiting battalions in USAREC, the model would have to include 39 dummy variables, or a separate predictive model would have to be calculated for each battalion. Lastly, a recruiter who decides to make recruiting a career cannot remain in the same unit for the remainder of his military service. If a recruiter who was successful in one unit is assigned to a unit where the model predicts failure, then should the recruiter be released? The answer would most likely be no, but this recruiter would be more likely to fail in his next unit. The PRiSM is used to help select candidate recruiters who do not depend on the unit to succeed and have the potential to be successful wherever they are assigned. While there is merit in investigating the effects of a unit's environment on recruiter success, the PRiSM should not depend on this environment.

C. RECOMMENDATIONS

This study was an initial phase of a more general study of the recruiter selection process. Further research is necessary, both to develop a more robust model, and to validate this initial model. The model's application in the recruiter selection process may have some effects on administrative policy. Some recommendations are offered to advance the study of the recruiter selection process and to improve the model developed in this paper.

- Redevelop the model using a larger sample size.
- Use the follow-up method in collecting data.
- Collect additional data not used or not available at the time of this study.
- Predictively validate PRiSM.
- Organize and reformat the USAREC information system to simplify analysis.
- Apply the PRiSM as an instrument to aid and improve recruiter selection.

1. Sample Size and the Follow-up Method of Testing

An issue of some concern was the limited size of the sample. It was the goal of the study to have a minimum of 500 recruiter records for the database, but circumstances made that goal infeasible. The limiting factor in the size of the database was the number of recruiters that could be tested with the SCT. Allocation of resources and the cooperation of USAREC can facilitate the expansion of the database. By using an expanded database and this methodology, the model should be redeveloped with the same candidate variables. A study that results in the same model, but with different-valued coefficients, attests to the efficacy of the variable selection for the PRiSM.

If the database were to be comprised of active recruiters, the SCT could be administered at each recruiting battalion's quarterly mission briefs. However, the disadvantages of conducting present-employee testing has already been discussed and should be avoided.

A more pragmatic approach to collecting data for the study is to administer a test or tests to soldiers just entering the ARS. This method is the follow-up method of testing. The test is first administered and the subject's performance is compared with the test scores periodically; at school, during TTE, and afterwards. The advantage of the follow-up method is that extraneous factors do not affect the soldier's performance on the test. A recruiter's experience in recruiting would not play a part in how well he does on the test.

Further, data for unsuccessful recruiters can be captured. These data were not available for this study, and may prove invaluable for further research. The reliability of the test for recruiters could also be measured by retesting the recruiters after a period of time and computing the correlation of the two scores. A high correlation between the scores would support the test's appropriateness for selection.

2. Additional Data Collection

A follow-up method would also allow the collection of other data not collected for this study. A growing concern in USAREC is the effect of stress on recruiter performance, causing discussion on the use of a stress test. A measure of stress could be included in the database, along with other factors which may be seen as indicators of success, but not defined in the list of successful recruiter traits in the current. Physical fitness scores, age, level of education, and other individual attributes could enhance the database.

3. Validation of the Prism

As an extension of the follow-up method for data collection, the model developed in this study can and should be validated. If the SCT had been administered to soldiers prior to entering the Army Recruiting School (ARS), a predicted success rate could be calculated for each student recruiter, using the PRiSM. At a minimum of nine months after successful completion of ARS, the actual performances of a recruiting class could be compared with their predicted performances. A correlation coefficient could then be computed for the group between these predicted and actual performances, and compared with the correlation coefficient of the PRiSM. A correlation which is close to the model's coefficient will support the validity of the model. A large number of students should be involved in the study to assure the stability of the correlation coefficient. Since USAREC holds 39 classes annually with an average of 40 students per class, validation of the PRiSM could be conducted within a two-year period.

The coefficients of the PRiSM should be recalibrated as more data become available. As more data are collected, the model can be recalculated using the same independent variables as in the final model of the current study. Because a follow-up method would be used, a greater range of data on all variables may result. Each iteration of validation could increase the accuracy of the model.

4. Organization of the USAREC Database

To facilitate analysis, the USAREC recruiter database must be organized for ease of access and accuracy. During the course of this research, considerable database manipulation was required to filter thousands of monthly records (belonging to over 400 reserve and active recruiters) into a coherent format. Many of the problems encountered during this process stemmed from lack of quality control of the database.

Names of recruiters were entered in different formats, sometimes for the same recruiter. One recruiter may be entered as last name, full first name, and middle initial, divided by commas. Another recruiter may be entered as last name, first initial, middle initial, no commas. Sometimes the initials would have periods, sometimes not. In combining data from different sources for one recruiter, his or her complete record may not appear. This happened on many occasions when a recruiter would be missing data for three to six months.

These recruiters' records became almost useless because of the discontinuity.

According to USAREC, input of data is required monthly for each of the recruiting battalions spread throughout the United States. The information is transmitted by modem to USAREC, at Fort Knox, Kentucky. However, some battalions are likely to enter information differently from others. USAREC must standardize information input and have information pass through one focal point in each battalion to ensure the correctness of the data.

Which elements to be included in the database are a subject for USAREC to decide. Further, USAREC must formally define what constitutes a successful recruiter. If defined in the strictest terms, and success is the accomplishment of monthly quotas, then monthly inputs would be needed. If USAREC defines success in aggregate terms, then quarterly or even annual inputs of missions assigned and achieved may satisfy database requirements. Factors of interest such as scores for SCT, AFQT, and educational data may become required inputs. In any event, USAREC must set the standards for its recruiter database construction.

5. The Taylor and Russell Tables

Incorporation of the PRiSM should occur when, or if, it is validated to the satisfaction of USAREC. The PRiSM should be used as a screening instrument to prevent selection of nonproductive recruiters. A useful tool to evaluate the usefulness of the PRiSM is the Taylor and Russell Tables, devised by H.C. Taylor and J.T. Russell for the Western Electric Company. [Ref. 25]

More formally called ***The Relationship of Validity Coefficients to the Practical Effectiveness of Tests in Selection: Discussion and Tables***, the Taylor and Russell Tables combine correlation coefficients with selection ratios and proportions of currently satisfactory selectees (base rates) to derive fractions of selected individuals who are expected to be successful. The selection ratio is the proportion of applicants who are selected. An excerpt from the Taylor and Russell Tables (Table 12) is shown below:

r	Selection Ratio				
	0.10	0.20	***0.30***	0.40	0.50
0.25	0.11	0.09	0.08	0.08	0.07
0.30	0.12	0.10	***0.09***	0.08	0.07
0.35	0.14	0.11	0.10	0.09	0.08
0.40	0.16	0.12	0.10	0.09	0.08

Table 12. Excerpt from Taylor and Russell Table (Base Rate = 0.05)

This portion of the tables is for a proportion of satisfactory applicants equal to **0.05**. The complete tables are in Appendix T. The employer has determined that the proportion of currently successful workers is **5%**. For instance, if the predictive tool has a correlation coefficient of **0.30** and the percent of applicants the employer wants to hire is **30%**, and the fraction of current workers deemed successful is 5%, then the fraction of those selected who are expected to be successful is **9%** . It can be seen that as the selection ratio decreases, the *expected* fraction of successful employees increases. If the proportion of currently satisfactory applicants varies, the Taylor and Russell Tables, which range from 0.05 to 0.95, can accommodate the change. In this manner USAREC can gauge the usefulness of the PRiSM. Chapter IV has presented a detailed application of the Taylor and Russell Tables.

When should the PRiSM be used to select recruiters? Timing is dependent on the administration of the SCT. When the SCT should be given to soldiers is beyond the scope of this paper, although several obvious opportunities exist. Since a recruiter must be a Noncommissioned Officer (NCO) in the rank of Sergeant (E-5) to Sergeant First Class (E-7), a recruiter must pass through several leadership " gates ". The first gate for an NCO is the Primary Leadership Development Course (PLDC). Selection to attend PLDC assumes that the soldier has attained the level of maturity and professionalism necessary to become a responsible NCO. Although there are exceptions, for the most part, this assumption is true. Another possible testing opportunity would be upon completion of basic training. However, most soldiers are young (18-21 years old), and changes in attitude and maturity level are predominant in this age group. It is highly likely that the scores for these soldiers would change dramatically over time and retesting would be required. Administration of the test and use of the PRiSM just prior to entrance to ARS is futile. The soldier has already been selected for recruiting and funds have been expended. Even when undesirable candidates are identified, administrative changes are time-consuming and monetary losses have already been borne by USAREC. PLDC seems to be the most advisable time for administering the SCT and computing the PRiSM results.

Once a predicted success rate has been computed for a soldier, it can become a part of his or her Military Personnel Records Jacket (MPRJ) and recorded on DA Form 201, which is a specific record of his military career. When primary branches need to fill a quota for recruiters, the PRiSM score can guide the assignment officer's nominations. Likewise, the RRP could examine the predicted success rates of soldiers for volunteer recruiters, and focus attention on those soldiers who meet or surpass the cut-off PRiSM score that USAREC establishes. As an end result, with potential failures screened out, the recruiters who enter the ARS have an increased likelihood of succeeding in recruiting duty.

APPENDIX A. INITIAL DATABASE FOR 276 RECRUITERS

This appendix includes the initial database in spread sheet format. Listed are the information obtained from USAREC, an administered sales ability test, and a personal questionnaire. Some of the column headings have been abbreviated, but are explained more clearly below:

- RECRUIT STATE: The current recruiting station the recruiter is assigned
- MONTH OS: Number of months a recruiter has been assigned to the recruiting station
- MONTHS MISSION: Number of months the recruiter was assigned a mission quota
- MONTHS ZERO: Number of months a recruiter achieved no enlistments
- MONTHS BOX: Number of months a recruiter achieved all assigned mission categories
- TOTAL GSA MISSION: Recruiter's cumulative number of assigned Graduate Senior Category "A" enlistment quotas
- TOTAL GSA ACH: Recruiter's cumulative number of achieved Graduate Senior Category "A" enlistment quotas
- TOTAL GSA PCT: Recruiter's percentage of achieved Graduate Senior Category "A" enlistment quotas
- TOTAL VOL MISSION: Recruiter's cumulative number of assigned other-than-Graduate Senior Category "A" enlistment quotas
- TOTAL VOL ACH: Recruiter's cumulative number of achieved other-than-Graduate Senior Category "A" enlistment quotas
- TOTAL VOL PCT: Recruiter's percentage of achieved other-than-Graduate Senior Category "A" enlistment quotas
- TEST: Recruiter raw score on the Sales Comprehension Test

- GSA PCT MON SUC: Percent of the months that a recruiter achieved GSA missions
- VOL PCT MON SUC: Percent of the months that a recruiter achieved VOL missions
- CUM ADJ PCT SUC: Summed weighted total of TOTAL GSA PCT and TOTAL VOL PCT (0.67 and 0.33, respectively)
- MONTHLY ADJ PCT SUC: Summed weighted total of GSA PCT MON SUC and VOL PCT MON SUC (0.67 and 0.33, respectively)

	RANK	PMOS	SELECTION	GENDER	MONTHS	HOME STATE	RECRUIT STATE	MONTHS OS	MONTHS MISSION	MONTHS ZERO
	MONTHS BOX	TOTAL GSA MSN	TOT GSA ACH	TOT GSA PCT	TOT VOL MSN	TOT VOL ACH	TOT VOL PCT	TEST		GSA PCT MON SUC
VOL PCT MON SUC	CUM ADJ PCT SUC	ACT MONTHS	MONTHLY ADJ PCT SUC							
	SSG	71G34	VOL	MALE	144	AR	CA	1	26	0
	5	14	38	271.43	31	82	264.52	29		1.00
1.00	269.15	143.00	100.00	1						
	SSG	51H34	VOL	MALE	148		MD	2	3	0
	0	1	2	200.00	1	2	200.00	1		1.00
1.00	200.00	146.00	100.00	1						
	SFC	OOR40	VOL	MALE	213	WA	CA	2	46	0
	13	11	14	127.27	16	17	106.25	32		0.00
1.00	120.34	211.00	33.00	0.33						
	SFC	OOR40	DA	MALE	108		MD	2	38	0
	8	10	3	30.00	10	10	100.00	38		
	53.10	106.00	0.00	0						
	SSG	88N30	VOL	MALE			MD	2	5	0
	0	1	3	300.00	2	4	200.00	6		1.00
1.00	267.00	-2.00	100.00	1						
	SFC	OOR40	VOL	MALE	216	SC	CA	2	48	0
	10	34	27	79.41	53	59	111.32	18		0.67
0.56	89.94	214.00	63.00	0.63000 03						
	SSG	OOR30	VOL	MALE	159	WV	CA	2	47	0
	17	42	28	66.67	75	72	96.00	4		0.40
0.73	76.35	157.00	50.80	0.50800 00						
	SSG	11B30	VOL	MALE	138		MD	3	2	0
	0	2	3	150.00	4	3	75.00	7		1.00
1.00	125.25	135.00	100.00	1						
	SSG	OOR30	DA	MALE	192		MD	3	36	0
	4	5	2	40.00	17	6	35.29	4		1.00
0.67	38.45	189.00	89.00	0.89000 01						
	SGT	71L20	DA	FEMALE	96		MD	3	6	0
	1	1	6	600.00	1	9	900.00	29		1.00
1.00	699.00	93.00	100.00	1						
	SGT	75B	VOL	MALE	142		MD	3	2	0
	0	2	1	50.00	2	2	100.00	63		0.50
1.00	66.50	139.00	66.50	0.665						
	SFC	OOR40	VOL	MALE	280		MD	3	23	0

	7	11	28	254.55	19	38	200.00	9		
	236.55	277.00	0.00	0						
	SFC	OOR40	DA	MALE	204	VA	CA	3	46	0
	12	44	37	84.09	76	74	97.37	39		0.50
0.50	88.47	201.00	50.00	0.5						
	SSG	12B30	DA	MALE	72		MD	3	4	0
	0	2	3	150.00	5	10	200.00	9		0.67
1.00	166.50	69.00	77.67	0.77666						
				68						
	SSG	94B34	VOL	FEMALE			MD	4	6	0
	0	1	1	100.00	1	2	200.00	20		1.00
1.00	133.00	-4.00	100.00	1						
	SSG	71G30	DA	FEMALE	105		MD	4	5	0
	0	3	1	33.33	6	4	66.67	45		0.67
0.67	44.33	101.00	66.67	0.66666						
				7						
	SSG	13B30	DA	MALE	132		MD	4	5	0
	0	3	3	100.00	4	9	225.00	25.5		0.75
1.00	141.25	128.00	83.25	0.8325						
	SSG	12B30	VOL	MALE	72		MD	4	6	0
	1	2	4	200.00	4	7	175.00	-11		1.00
1.00	191.75	68.00	100.00	1						
	SFC	OOR40	VOL	MALE	168	GA	CA	4	44	0
	13	5	15	300.00	15	43	286.67	-6		0.88
0.50	295.60	164.00	75.13	0.75125						
	SSG	31U34	VOL	MALE	107	NM	CA	4	6	0
	0	3	3	100.00	6	5	83.33	-32		0.60
1.00	94.50	103.00	73.20	0.732						
	SSG	93C30	DA	MALE			MD	4	6	0
	1	1	4	400.00	4	4	100.00	48		0.75
1.00	301.00	-4.00	83.25	0.8325						
	SSG	OOR30	DA	MALE	138	PA	CA	4	45	0
	7	38	22	57.89	67	55	82.09	33		0.40
0.56	65.88	134.00	45.13	0.45133						
				34						
	SSG	E34	DA	MALE	178	PR	CA	4	5	0
	0	2	2	100.00	2	5	250.00	47		0.60
1.00	149.50	174.00	73.20	0.732						
	SSG	11B30	VOL	MALE	122		MD	4	6	0
	1	3	4	133.33	6	6	100.00	27		1.00
0.75	122.33	118.00	91.75	0.9175						
	SSG	51T34	DA	MALE	108	IL	CA	5	4	0
	0	1	2	200.00	2	3	150.00	20		1.00
0.60	183.50	103.00	86.80	0.868						
	SSG	91B34	DA	MALE	84	NC	CA	5	7	0
	0	5	6	120.00	8	8	100.00	11		0.50
0.67	113.40	79.00	55.50	0.55500						
				01						
	SSG	OOR30	DA	MALE	120	CA	CA	5	50	0
	10	38	39	102.63	67	76	113.43	26		0.27
0.57	106.20	115.00	37.13	0.37129						
				86						
	SSG	11M34	VOL	MALE	132	NC	CA	5	6	0
	1	2	2	100.00	5	5	100.00	53		0.80
0.60	100.00	127.00	73.40	0.734						
	SSG	92A34	DA	MALE	99		MD	5	6	0
	1	2	4	200.00	5	8	160.00	29		0.75
0.75	186.80	94.00	75.00	0.75						
	SSG	74F34	VOL	FEMALE	72		MD	5	7	0
	2	2	8	400.00	5	13	260.00	17		1.00
1.00	353.80	67.00	100.00	1						
	SSG	67V34	DA	MALE	150		MD	5	7	0
	1	3	8	266.67	7	10	142.86	25		1.00
1.00	225.81	145.00	100.00	1						
	SGT	74F24	DA	MALE	111		MD	6	7	0

	0	3	2	66.67	6	7	116.67	51		1.00
1.00	83.17	105.00	100.00	1						
	SFC	11M44	DA	MALE	136		MD	6	8	0
	0	3	4	133.33	6	7	116.67	7		1.00
1.00	127.83	130.00	100.00	1						
	SSG	55B34	DA	MALE	119	NY	CA	6	13	0
	2	5	5	100.00	6	6	100.00	26		0.50
0.83	100.00	113.00	61.00	0.60999						
				99						
	SGT	19K24	DA	MALE	84		MD	6	8	0
	1	2	4	200.00	7	7	100.00	-24		1.00
1.00	167.00	78.00	100.00	1						
	SFC	OOR40	DA	MALE	210		MD	6	39	0
	7	3	3	100.00	8	10	125.00	34		1.00
0.00	108.25	204.00	67.00	0.67						
	SSG	76J34	DA	MALE	93		MD	7	10	0
	0	4	4	100.00	9	16	177.78	14		1.00
1.00	125.67	86.00	100.00	1						
	SFC	OOR40	DA	MALE	192	MI	CA	7	28	0
	5	8	10	125.00	23	43	186.96	-7		1.00
1.00	145.45	185.00	100.00	1						
	SFC	OOR40	VOL	MALE	168		MD	7	8	0
	1	5	7	140.00	11	17	154.55	14		0.50
1.00	144.80	161.00	66.50	0.665						
	SSG	88H34	DA	MALE	120		MD	7	10	0
	3	4	5	125.00	9	15	166.67	40		1.00
1.00	138.75	113.00	100.00	1						
	SFC	OOR40	DA	MALE	216	HI	CA	8	45	0
	15	33	23	69.70	73	72	98.63	27		0.67
0.83	79.24	208.00	72.17	0.72166						
				65						
	SSG	11B34	DA	MALE	132	CA	CA	8	11	0
	0	5	6	120.00	12	7	58.33	-2		1.00
0.50	99.65	124.00	83.50	0.835						
	SFC	45K44	DA	MALE	156		MD	8	11	0
	2	3	10	333.33	9	25	277.78	37		1.00
0.00	315.00	148.00	67.00	0.67						
	SSG	13B34	DA	MALE	120	CA	CA	8	10	0
	2	7	6	85.71	9	14	155.56	21		1.00
1.00	108.76	112.00	100.00	1						
	SGT	71L24	DS	FEMALE	84		MD	8	15	0
	0	4	8	200.00	11	13	118.18	9		0.50
1.00	173.00	76.00	66.50	0.665						
	SSG	12B30	VOL	MALE	144		MD	8	11	0
	1	2	1	50.00	3	6	200.00	-27		1.00
1.00	99.50	136.00	100.00	1						
	SSG	11B34	DA	MALE	113	NC	CA	8	9	0
	0	5	5	100.00	11	13	118.18	18		1.00
0.00	106.00	105.00	67.00	0.67						
	SSG	OOR30	DA	MALE			MD	8	12	0
	2	6	7	116.67	10	16	160.00	31.5		0.50
1.00	130.97	-8.00	66.50	0.665						
	SSG	19K34	DA	MALE	128	MD	CA	9	10	0
	2	4	7	175.00	8	16	200.00	33		1.00
1.00	183.25	119.00	100.00	1						
	SSG	11H34	DA	MALE	81	KY	CA	9	10	0
	0	6	5	83.33	12	12	100.00	60		0.67
0.75	88.83	72.00	69.42	0.69416						
				68						
	SGT	67U24	DA	MALE	78	OK	CA	9	11	0
	0	6	11	183.33	9	17	188.89	1		1.00
1.00	185.17	69.00	100.00	1						
	SFC	54B44	DA	MALE	160		MD	9	12	0
	1	6	5	83.33	12	16	133.33	-11		0.50
1.00	99.83	151.00	66.50	0.665						

	SSG	11B34	DA	MALE	111	GA	CA	9	12	0
	0	11	5	45.45	17	14	82.35	1		0.50
1.00	57.63	102.00	66.50	0.665						
	SFC	OOR40	VOL	MALE	199	HI	CA	9	49	0
	13	37	21	56.76	60	45	75.00	22		0.20
0.50	62.78	190.00	29.90	0.299						
	SSG	11B34	DA	MALE	36		MD	9	11	0
	3	7	19	271.43	12	32	266.67	64		0.75
0.50	269.86	27.00	66.75	0.6675						
	SSG	14S34	DA	MALE	79	FL	CA	9	12	0
	2	9	14	155.56	13	21	161.54	-38		1.00
0.50	157.53	70.00	83.50	0.835						
	SGT	88M24	DA	MALE	95	CA	CA	9	9	0
	0	5	8	160.00	10	11	110.00	11		1.00
0.00	143.50	86.00	67.00	0.67						
	SGT	71L24	DA	MALE	135		MD	9	11	0
	3	5	7	140.00	10	14	140.00	34		
	SSG	63B34	DA	MALE	144		MD	9	11	0
	1	7	7	100.00	13	12	92.31	10		0.75
0.50	97.46	135.00	66.75	0.6675						
	SGT	71G24	DA	MALE	96		MD	9	13	0
	3	6	4	66.67	14	22	157.14	0		1.00
0.50	96.52	87.00	83.50	0.835						
	SSG	11B34	DA	MALE	36	IN	CA	9	10	0
	0	4	2	50.00	15	9	60.00	10		0.67
0.33	53.30	27.00	55.67	0.55666 67						
	SSG	63B34	DA	MALE	130	MS	CA	9	12	0
	0	5	9	180.00	13	13	100.00	28		
	153.60	121.00		0						
	SGT	77F24	DA	MALE	107		MD	9	11	0
	2	5	8	160.00	9	18	200.00	5		1.00
1.00	173.20	98.00	100.00	1						
	SSG	91M34	DA	FEMALE	136		MD	9	10	0
	1	5	5	100.00	11	11	100.00	30		0.50
0.50	100.00	127.00	50.00	0.5						
	SGT	92A20	DA	MALE	144	PAGO PAGO	CA	10	10	0
	0	2	2	100.00	3	3	100.00	-1		0.75
0.75	100.00	134.00	75.00	0.75						
	SSG	63B34	DA	MALE	122		MD	10	12	0
	3	5	7	140.00	11	20	181.82	2		0.50
1.00	153.80	112.00	66.50	0.665						
	SSG	67T34	VOL	MALE	119	CA	CA	10	13	0
	4	11	12	109.09	18	19	105.56	4		0.80
0.80	107.92	109.00	80.00	0.8						
	SSG	11B34	DA	MALE	127	WA	CA	10	14	0
	3	9	12	133.33	18	18	100.00	23		1.00
0.25	122.33	117.00	75.25	0.7525						
	SSG	13F34	VOL	MALE	102	OH	CA	10	13	0
	2	6	11	183.33	18	19	105.56	17		0.40
0.50	157.67	92.00	43.30	0.433						
	SFC	OOR40	DA	MALE	243	NM	CA	10	48	0
	12	26	9	34.62	43	19	44.19	21		0.60
0.29	37.77	233.00	49.63	0.49628 56						
	SSG	63B34	DA	MALE	156	CA	CA	10	11	0
	3	9	10	111.11	14	22	157.14	32		1.00
0.80	126.30	146.00	93.40	0.934						
	SGT	71L24	VOL	MALE	84	MO	CA	10	11	0
	3	7	6	85.71	16	12	75.00	-1		1.00
0.33	82.18	74.00	78.00	0.77999 98						
	SSG	91T34	DA	FEMALE	142	CA	CA	10	13	0
	0	5	5	100.00	10	6	60.00	4		0.75

0.00	86.80	132.00	50.25	0.5025						
	SSG	13B34	DA	MALE	102	CA	CA	10	12	0
	2	8	10	125.00	14	23	164.29	19		1.00
0.50	137.96	92.00	83.50	0.835						
	SGT	76V24	DA	MALE	156	CA	CA	10	12	0
	3	8	10	125.00	18	22	122.22	41		0.75
0.80	124.08	146.00	76.65	0.7665						
	SFC	OOR40	DA	MALE	204		MD	10	40	0
	12	30	30	100.00	60	62	103.33	33		0.75
0.20	101.10	194.00	56.85	0.5685						
	SSG	11M34	DA	MALE	118	OK	CA	11	14	0
	2	11	9	81.82	18	11	61.11	0		0.40
1.00	74.98	107.00	59.80	0.598						
	SSG	13B34	DA	MALE	86		MD	11	13	0
	4	10	9	90.00	18	24	133.33	17		0.67
1.00	104.30	75.00	77.67	0.77666						
				68						
	SSG	75B34	VOL	MALE	96		MD	11	13	0
	2	6	12	200.00	15	18	120.00	7		0.83
0.33	173.60	85.00	66.83	0.66833						
				3						
	SFC	63N44	DA	MALE	142	SD	CA	11	14	0
	3	9	10	111.11	17	20	117.65	33		
	SSG	95B34	DA	MALE	120		MD	12	15	0
	4	11	16	145.45	20	28	140.00	27		0.75
1.00	143.65	108.00	83.25	0.8325						
	SGT	E25	DA	MALE	104	CA	CA	12	15	0
	2	10	12	120.00	16	21	131.25	3		0.63
1.00	123.71	92.00	74.88	0.74875						
	SGT	93F24	DA	FEMALE	86		MD	12	15	0
	4	5	9	180.00	17	19	111.76	35		0.83
0.50	157.48	74.00	72.33	0.72333						
				31						
	SSG	91B34	VOL	MALE	231	GU	CA	13	16	0
	9	15	24	160.00	23	41	178.26	33		1.00
0.88	166.03	218.00	95.88	0.95875						
	SSG	11M34	DA	MALE	115	TX	CA	13	16	0
	4	12	8	66.67	18	19	105.56	10		0.71
0.67	79.50	102.00	69.86	0.69857						
				17						
	SGT	19D24	VOL	MALE	88	CA	CA	13	16	0
	4	13	9	69.23	21	20	95.24	-20		0.57
0.67	77.81	75.00	60.29	0.60285						
				75						
	SSG	31C34	DA	MALE	144	WI	CA	13	16	0
	2	12	7	58.33	20	17	85.00	39		0.67
0.50	67.13	131.00	61.17	0.61166						
				68						
	SSG	24T34	DA	MALE	127	CA	CA	13	16	0
	1	13	5	38.46	23	17	73.91	25		0.38
1.00	50.16	114.00	58.13	0.58125						
	SSG	E34	DA	MALE	48	CA	CA	13	16	0
	3	11	11	100.00	17	20	117.65	36		0.75
0.75	105.82	35.00	75.00	0.75						
	SGT	92Y20	DA	MALE	105		MD	13	16	0
	5	7	12	171.43	19	23	121.05	22		0.88
0.71	154.80	92.00	82.20	0.82196						
				43						
	SSG	19D34	DA	MALE	96	FL	CA	13	15	0
	4	12	11	91.67	22	22	100.00	28		0.56
0.67	94.42	83.00	59.22	0.59222						
				26						
	SSG	63H34	DA	MALE	120	MI	CA	13	16	0
	2	16	10	62.50	31	26	83.87	21		0.50
0.63	69.55	107.00	54.13	0.54125						

	SSG	E34	DA	FEMALE	122	NY	CA	13	15	0
	4	9	17	188.89	18	27	150.00	26		0.86
0.67	176.06	109.00	79.43	0.79428						
				59						
	SSG	95B34	DA	MALE			MD	13	15	0
	5	10	14	140.00	21	29	138.10	44		0.75
0.57	139.37	-13.00	69.11	0.69107						
				15						
	SSG	95B34	VOL	MALE	162		MD	13	15	0
	3	13	10	76.92	19	16	84.21	23		0.63
0.50	79.33	149.00	58.38	0.58375						
	SSG	11B34	VOL	MALE	156	WY	CA	14	9	0
	0	8	3	37.50	13	14	107.69	44		0.50
0.50	60.66	142.00	50.00	0.5						
	SFC	73C44	DA	MALE	157	CA	CA	14	16	0
	3	11	11	100.00	24	24	100.00	31		0.78
0.75	100.00	143.00	76.86	0.76861						
				12						
	SSG	95B34	DA	MALE	96		MD	14	16	0
	3	14	16	114.29	23	35	152.17	18		0.89
0.50	126.79	82.00	76.06	0.76055						
				56						
	SSG	62B34	DA	MALE	120	MT	CA	14	16	0
	2	16	12	75.00	24	25	104.17	33		0.63
0.90	84.63	106.00	71.58	0.71575						
	SSG	29J34	DA	MALE	172	GU	CA	14	15	0
	2	11	9	81.82	17	15	88.24	-20		0.75
0.67	83.94	158.00	72.25	0.72250						
				01						
	SSG	31L34	DA	MALE	108	CA	CA	14	17	0
	2	9	11	122.22	18	24	133.33	38		0.75
0.67	125.89	94.00	72.25	0.72250						
				01						
	SSG	91M34	DA	MALE	153	ID	CA	14	17	0
	6	15	21	140.00	25	26	104.00	44		0.89
0.50	128.12	139.00	76.06	0.76055						
				56						
	SSG	71G34	DA	MALE	74	MI	CA	14	16	0
	1	17	8	47.06	34	24	70.59	6		0.56
0.43	54.82	60.00	51.37	0.51365						
				09						
	SSG	88H34	DA	MALE	96	OH	CA	14	16	0
	4	8	14	175.00	17	27	158.82	6		0.88
0.71	169.66	82.00	82.20	0.82196						
				43						
	SSG	95B34	DA	MALE	132	ND	CA	14	17	0
	3	14	12	85.71	23	21	91.30	3		0.63
0.67	87.56	118.00	63.88	0.63875						
				01						
	SSG	11B34	VOL	MALE	103		MD	14	16	0
	4	13	12	92.31	22	23	104.55	29		0.88
0.80	96.35	89.00	85.03	0.85025						
	SSG	88M34	DA	MALE	98	HI	CA	14	17	0
	4	10	17	170.00	19	27	142.11	35		0.88
0.71	160.79	84.00	82.20	0.82196						
				43						
	SGT	88M24	DA	MALE	84		MD	15	16	0
	5	12	12	100.00	21	23	109.52	14	Correa,	0.60
0.75	103.14	69.00	64.95	0.6495						
	SSG	13B34	DS	MALE	84		MD	15	19	0
	3	9	10	111.11	16	17	106.25	-4	Jones,	0.83
0.60	109.51	69.00	75.63	0.75633						
				31						
	SSG	75B34	DA	MALE	144		MD	15	16	0
	215.88	129.00	60.91	0.60909						

				09						
	SSG	88M34	DA	MALE	132		MD	15	17	0
	3	9	9	100.00	22	19	86.36	23		0.80
0.50	95.50	117.00	70.10	0.701						
	SFC	OOR40	DA	MALE	240		MD	16	38	0
	7	8	8	100.00	9	10	111.11	-13		0.67
1.00	103.67	224.00	77.67	0.77666						
				68						
	SFC	·71L44	DA	MALE	165	IN	CA	16	18	0
	2	11	8	72.73	21	25	119.05	42		0.45
0.86	88.01	149.00	58.74	0.58740						
				23						
	SSG	OOR30	DA	MALE	48		MD	16	20	0
	4	15	13	86.67	25	33	132.00	49		0.64
0.82	101.63	32.00	69.64	0.69636						
				39						
	SSG	51H34	DA	MALE	132	WI	CA	16	20	0
	5	14	18	128.57	25	32	128.00	27		1.00
0.50	128.38	116.00	83.50	0.835						
	SSG	13R34	DA	MALE	122		MD	16	18	0
	2	13	15	115.38	25	28	112.00	56		0.82
0.50	114.27	106.00	71.32	0.71318						
				19						
	SSG	OOR30	VOL	MALE	156		MD	16	38	0
	13	5	7	140.00	10	13	130.00	-5		0.63
0.83	136.70	140.00	69.37	0.69374						
				98						
	SFC	OOR40	VOL	MALE	198	OH	CA	17	39	0
	7	2	2	100.00	12	5	41.67	3		1.00
0.67	80.75	181.00	89.00	0.89000						
				01						
	SSG	13B34	DA	MALE	161		MD	17	19	0
	6	15	15	100.00	27	36	133.33	27		0.45
0.60	111.00	144.00	50.25	0.50254						
				51						
	SSG	13P34	DA	MALE	123		MD	17	21	0
	3	13	12	92.31	27	31	114.81	-18		0.60
0.78	99.74	106.00	65.87	0.65866						
				67						
	SSG	91B34	DA	MALE	144		MD	17	19	0
	5	12	15	125.00	28	36	128.57	3		0.77
0.82	126.18	127.00	78.54	0.78538						
				48						
	SSG	11M34	DA	MALE	144		MD	17	19	0
	4	14	15	107.14	23	29	126.09	23.5	Vete,	0.85
0.88	113.39	127.00	85.57	0.85567						
				31						
	SSG	13R34	DA	MALE	150	TN	CA	18	18	0
	5	17	13	76.47	29	35	120.69	20		0.58
0.75	91.06	132.00	63.83	0.63833						
				31						
	SSG	13M34	DA	MALE	94	AL	CA	18	20	0
	5	15	18	120.00	27	34	125.93	36		0.50
0.86	121.96	76.00	61.79	0.61785						
				71						
	SGT	31V24	DA	MALE	121		MD	18	19	0
	4	11	17	154.55	24	31	129 17	24		0.67
0.50	146 17	103.00	61.17	0 61166						
				68						
	SSG	11B34	DA	MALE	12		MD	18	19	0
	5	10	13	130.00	25	43	172.00	8		0.75
0.67	143.86	-6.00	72.25	0.72250						
				01						
	SSG	93C34	DA	MALE	144	IL	CA	18	19	0
	2	15	8	53.33	31	17	54.84	31		0.58

0.50	53.83	126.00	55.58	0.55583						
				33						
	SFC	OOR40	VOL	MALE	168		MD	18	40	0
	10	20	9	45.00	29	24	82.76	-19		0.27
0.60	57.46	150.00	38.07	0.38072						
				70						
	SGT	74F24	DA	MALE	88		MD	18	20	0
	2	16	10	62.50	27	31	114.81	23		0.50
0.77	79.76	70.00	58.88	0.58884						
				62						
	SSG	11B34	DA	MALE	128	CA	CA	18	20	0
	5	17	19	111.76	28	28	100.00	32		0.92
0.67	107.88	110.00	83.85	0.83846						
				17						
	SGT	45B24	DA	MALE	122	AZ	CA	18	19	0
	2	24	17	70.83	33	33		28		0.67
0.67	80.46	104.00	66.67	0.66666						
				7						
	SFC	82C34	DA	MALE	125	CO	CA	18	20	0
	4	22	23	104.55	36	49	136.11	[illegible]		0.92
0.75	114.96	107.00	86.60	0.86596						
				15						
	SGT	92Y24	DA	MALE	108		MD	18	20	0
	4	14	7	50.00	28	39	139.29	17		.79
0.85	79.46	90.00	80.57	0.80565						
				92						
	SSG	13M34	VOL	MALE	150	SD	CA	18	20	0
	5	15	14	93.33	22	40	181.82	20		0.70
1.00	122.53	132.00	79.90	0.799						
	SSG	13R34	DA	MALE	148	AZ	CA	19	21	0
	4	16	7	43.75	32	27	84.38	0		0.46
0.90	57.16	129.00	60.62	0.60623						
				04						
	SSG	95B34	DA	MALE	120		MD	19	19	0
	7	16	22	137.50	24	35	145.83	40		0.77
0.56	140.25	101.00	69.87	0.69871						
				82						
	SSG	E35	DA	MALE	150	TX	CA	19	20	0
	5	18	21	116.67	32	38	118.75	25		0.79
0.44	117.35	131.00	67.31	0.67309						
				49						
	SSG	13P34	DA	MALE	120	TX	CA	19	23	0
	2	19	14	73.68	31	39	125.81	31		0.20
0.89	90.88	101.00	42.73	0.42733						
				33						
	SSG	91C34	DA	MALE	63		MD	19	20	0
	5	12	8	66.67	25	27	108.00	24		0.62
0.60	80.31	44.00	61.03	0.61030						
				79						
	SSG	76Y34	DS	MALE	84		MD	20	23	0
	7	15	24	160.00	27	45	166.67	34		0.77
0.67	162.20	64.00	73.54	0.73538						
				47						
	SSG	11B34	DA	MALE	161	FL	CA	20	22	0
	5	16	24	150.00	28	43	153.57	45		0.82
0.63	151.18	141.00	75.44	0.75443						
				19						
	SSG	OOR30	VOL	MALE			MD	20	19	0
	4	15	7	46.67	26	27	103.85	28		0.62
0.78	65.54	-20.00	66.90	0.66897						
				46						
	SSG	11C34	DA	MALE	185	MO	CA	20	22	0
	6	16	12	75.00	30	21	70.00	20		0.64
0.50	73.35	165.00	59.57	0.59571						
				41						

	SGT	92A20	VOL	MALE	96		MD	20	22	0
	5	12	20	166.67	31	41	132.26	10		0.85
0.55	155.31	76.00	74.69	0.74692						
				33						
	SGT	95B24	DA	MALE	86		MD	20	22	0
	4	16	19	118.75	32	49	153.13	26		0.73
0.85	130.09	66.00	77.06	0.77056						
				39						
	SSG	88M34	DA	MALE	108	NC	CA	20	25	0
	6	16	19	118.75	30	60	200.00	26		0.69
0.77	145.56	88.00	71.77	0.71769						
				25						
	SFC	OOR40	DA	MALE	156		MD	20	40	0
	10	35	35	100.00	65	58	89.23	-26		0.71
0.36	96.45	136.00	59.86	0.59857						
				15						
	SGT	74F24	DA	MALE			MD	21	23	0
	7	19	29	152.63	34	49	144.12	30		0.87
0.58	149.82	-21.00	77.32	0.77316						
				67						
	SGT	92A20	DA	MALE	87		MD	21	24	0
	3	15	12	80.00	32	32	100.00	57		1.00
0.50	86.60	66.00	83.50	0.835						
	SSG	76Y34	DA	MALE	162	TN	CA	21	24	0
	3	14	19	135.71	23	25	108.70	37		0.73
0.38	126.80	141.00	61.10	0.61102						
				29						
	SSG	92A30	DA	MALE	118		MD	21	24	0
	4	19	17	89.47	39	46	117.95	47		0.60
0.71	98.87	97.00	63.77	0.63771						
				43						
	SSG	31Q34	DA	MALE	165	CA	CA	22	24	0
	8	16	32	200.00	31	51	164.52	14		0.82
0.45	188.29	143.00	70.18	0.70176						
				42						
	SFC	71L44	DA	FEMALE	162		MD	22	26	0
	6	17	14	82.35	32	37	115.63	31		0.69
0.69	93.33	140.00	69.23	0.69230						
				8						
	SGT	74D24	DA	FEMALE		MI	CA	22	23	0
	4	13	16	123.08	32	39	121.88	17		0.50
0.90	122.68	-22.00	63.20	0.632						
	SSG	19K34	DA	MALE		NY	CA	22	22	0
	7	16	21	131.25	33	45	136.36	-3		0.71
0.77	132.94	-22.00	72.68	0.72678						
				71						
	SFC	OOR40	DA	MALE	209	CA	CA	22	48	0
	15	44	43	97.73	71	83	116.90	32		0.67
0.60	104.05	187.00	64.47	0.64466						
				68						
	SSG	91C34	DA	FEMALE	166		MD	22	24	0
	3	20	22	110.00	38	47	123.68	5		0.50
0.58	114.52	144.00	52.75	0.52749						
				98						
	SSG	19K34	DA	MALE		WV	CA	22	24	0
	6	26	20	76.92	36	34	94.44	29		0.56
0.63	82.71	-22.00	58.31	0.58312						
				5						
	SSG	27F34	DA	MALE	119	KY	CA	22	22	0
	6	17	13	76.47	27	31	114.81	26		0.47
0.77	89.12	97.00	56.65	0.56651						
				31						
	SFC	OOR40	DA	MALE	216		MD	23	37	0
	7	7	8	114.29	13	19	146.15	16		0.50
0.33	124.80	193.00	44.50	0.44499						

				98						
	SSG	96R34	VOL	MALE	156		MD	23	23	0
	8	20	26	130.00	34	54	158.82	42		0.92
0.78	139.51	133.00	87.08	0.8708[illegible]						
				36						
	SSG	11B34	VOL	MALE	12C	NM	CA	23	24	0
	7	26	21	80.77	39	36	92.31	-7		0.50
0.75	84.58	97.00	58.25	0.5825						
	SSG	74C34	VOL	MALE	119	NC	CA	23	25	0
	8	12	35	291.67	26	55	211.54	-1		1.00
0.63	265.22	96.00	87.63	0.87625						
	SSG	19K34	DA	MALE	96	CA	CA	23	24	0
	8	25	23	92.00	37	38	102.70	48	Duda,	0.65
0.64	95.53	73.00	64.35	0.64352						
				96						
	SSG	11M34	VOL	MALE	98	WA	CA	24	28	0
	5	23	18	78.26	38	49	128.95	14		0.47
0.82	94.99	74.00	58.27	0.58266						
				69						
	SGT	12B24	DA	MALE	93		MD	24	25	0
	6	15	17	113.33	30	45	150.00	33		0.63
0.92	125.43	69.00	72.34	0.72336						
				54						
	SSG	13M34	DA	MALE	81	CA	CA	24	28	0
	4	22	13	59.09	38	25	65.79	24		0.44
0.46	61.30	57.00	44.54	0.44543						
				25						
	SFC	OOR40	DA	MALE	176	CA	CA	24	50	0
	17	40	63	157.50	73	93	127.40	32		0.83
0.75	147.57	152.00	80.26	0.80264						
				25						
	SSG	92Y34	DS	MALE	108		MD	24	28	0
	6	23	21	91.30	37	32	86.49	-3		0.63
0.50	89.71	84 00	58.38	0.58375						
	SSG	95B34	DA	MALE	149		MD	25	27	0
	9	23	37	160.87	39	54	138.46	48		1.00
0.60	153.47	124.00	86.80	0.868						
	SSG	19D34	DA	MALE	148	CA	CA	25	25	0
	5	23	16	69.57	43	46	106.98	21		0.65
0.81	81.91	123.00	70.36	0.70362						
				5						
	SFC	OOR40	VOL	MALE	180		MD	25	40	0
	15	41	39	95.12	70	74	105.71	23		0.65
0.72	98.62	155.00	67.52	0.67521						
				25						
	SSG	95B34	DA	MALE			CA	25	28	0
	5	21	19	90.48	39	29	74.36	0		0.55
0.50	85.16	-25.00	53.35	0.5335						
	SSG	54B34	DA	MALE	144	SC	CA	25	28	0
0.58	0.82	115 81	119.00	65.79	0.65789					
					45					
	SFC	OOR40	DA	MALE	144		MD	25	38	0
	14	29	39	134.48	51	62	121.57	31	p,	0.76
0.67	130.22	119 00	72 83	0.72827						
				61						
	SSG	19K34	VOL	MALE	145		MD	25	27	0
	6	21	29	138.10	39	53	135.90	56.5		0 62
0.50	137.37	120 00	57 98	0.57976						
				21						
	SFC	54B44	DA	MALE	165		MD	25	27	0
	7	19	42	221.05	38	69	181.58	17		0.85
0.69	208.03	140 00	79 64	0 79637						
				5						
	SGT	11H24	DA	MALE	120		MD	26	28	0
	8	25	24	96.00	42	39	92.86	4		0.73

0.53	94.96	94.00	66.33	0.66327						
				27						
	SSG	19D34	DA	MALE	144	FL	CA	26	28	0
	12	24	21	87.50	42	49	116.67	31		0.71
0.69	97.13	118.00	70.70	0.70703						
				32						
	SSG	12B34	DA	MALE	132	NC	CA	26	29	0
	9	25	20	80.00	43	42	97.67	18		0.81
0.57	85.83	106.00	73.10	0.73095						
				26						
	SSG	13B34	DA	MALE	138	GU	CA	26	26	0
	7	29	29	100.00	47	50	106.38	32		0.65
0.69	102.11	112.00	66.24	0.66237						
				5						
	SFC	95B44	DA	MALE	140	MO	CA	26	27	0
	15	23	26	113.04	39	60	153.85	0		0.95
1.00	126.51	114.00	96.47	0.96473						
				65						
	SGT	19D24	DS	MALE	156		MD	26	31	0
	5	18	23	127.78	33	38	115.15	26		0.92
0.56	123.61	130.00	79.75	0.79750						
				03						
	SSG	19D34	DA	MALE	141	ID	CA	26	26	0
	2	15	12	80.00	31	24	77.42	20		0.57
0.50	79.15	115.00	54.79	0.54785						
				74						
	SSG	76P34	DA	MALE	126	AL	CA	26	28	0
	6	26	22	84.62	46	47	102.17	-23		0.58
0.75	90.41	100.00	63.54	0.63539						
				44						
	SFC	95B44	DA	MALE	158	OK	CA	26	25	0
	5	14	15	107.14	31	35	112.90	47		0.72
0.56	109.04	132.00	66.95	0.66951						
				37						
	SSG	95B34	DA	MALE	136		MD	26	27	0
	8	27	21	77.78	46	48	104.35	39		0.56
0.71	86.55	110.00	60.79	0.60793						
				69						
	SFC	54B34	DA	MALE	180	OR	CA	26	28	0
	8	29	22	75.86	43	40	93.02	-6	0.68	0.45
81.53	154.00	60.84	0.6084212							
	SSG	13B34	DA	MALE	98	MO	CA	26	29	0
	5	23	26	113.04	37	37	100.00	50		0.52
0.69	108.74	72.00	57.94	0.57941						
				43						
	SSG	19K34	DA	MALE	153	MI	CA	26	27	0
	5	24	16	66.67	37	34	91.89	20		0.72
0.36	74.99	127.00	60.39	0.60388						
				86						
	SSG	19D34	DA	MALE	108	CA	CA	27	28	0
	5	28	23	82.14	42	51	121.43	35		0.78
1.00	95.11	81.00	85.11	0.85111						
				11						
	SFC	11B44	DA	MALE	163	IA	CA	27	30	0
	7	24	20	83.33	35	42	120.00	17	Kasal,	0.86
1.00	95.43	136.00	90.43	0.90428						
				58						
	SFC	11M44	DA	MALE	180		MD	27	29	0
	8	23	28	121.74	42	55	130.95	24		0.67
0.56	124.78	153.00	63.00	0.63000						
				03						
	SFC	77F44	DA	MALE	156		MD	28	29	0
	8	27	24	88.89	46	49	106.52	3		0.75
0.50	94.71	128.00	66.75	0.6675						
	SSG	13R34	DA	MALE	182	IL	CA	29	27	0

0.82	7 133.37	18 153.00	22 71.67	122.22 0.71666 69	25	39	156.00	35		0.67
0.75	SGT 9 83.13	11C24 32 84.00	DA 26 63.83	MALE 81.25 0.63833 31	113 46	KY 40	CA 86.96	29 -15	31	0 0.58
0.63	SGT 11 101.83	13B24 24 91.00	DA 26 59.71	MALE 108.33 0.59708 31	120 44	SC 39	CA 88.64	29 21	32	0 0.58
0.86	SFC 14 132.57	67S44 30 119.00	DA 39 81.89	MALE 130.00 0.81885 71	148 45	NY 62	CA 137.78	29 36	32	0 0.80
0.78	SFC 15 161.29	91D44 26 140.00	DA 39 75.92	MALE 150.00 0.75916 67	169 38	70	MD 184.21	29 53	32	0 0.75
0.62	SSG 12 102.53	11H34 32 51.00	DA 31 64.97	MALE 96.88 0.64974 39	80 50	CO 57	CA 114.00	29 38	31	0 0.67
0.55	SSG 6 96.21	11H34 22 96.00	DS 20 51.50	MALE 90.91 0.51500 01	125 43	46	MD 106.98	29 15	31	0 0.50
0.75	SGT 10 106.27	31L24 34 79.00	DA 35 71.36	MALE 102.94 0.71358 68	108 46	IL 52	CA 113.04	29 20	31 k,	0 0.70
0.70	SFC 7 77.87	95B44 32 132.00	DA 25 67.77	MALE 78.13 0.67766 68	162 53	IL 41	CA 77.36	30 36	32	0 0.67
0.50	SGT 10 81.82	95B24 27 78.00	DA 23 56.70	MALE 85.19 0.567	108 44	33	MD 75.00	30 27	32	0 0.60
0.33	SGT 8 106.98	19K24 28 114.00	DA 33 55.67	MALE 117.86 0.55666 67	144 53	45	MD 84.91	30 12	32	0 0.67
0.62	SSG 10 108.75	16P34 29 -30.00	DA 31 62.94	MALE 106.90 0.62944 09	48	MI 54	CA 112.50	30 9	34	0 0.64
0.43	SSG 5 66.24	16S34 28 138.00	DA 18 52.01	MALE 64.29 0.52012 38	168 47	KY 33	CA 70.21	30 49	31	0 0.57
0.60	SFC 11 133.87	OOR40 10 104.00	DA 14 58.88	MALE 140.00 0.58883 31	135 14	17	MD 121.43	31 31	38	0 0.58
0.67	SSG 8 105.46	13B34 28 108.00	DA 29 66.67	MALE 103.57 0.66666 7	139 43	TX 47	CA 109.30	31 26	29	0 0.67
0.93	SSG 7 93.77	19D34 29 96.00	VOL 26 70.84	MALE 89.66 0.70842 84	127 47	IL 48	CA 102.13	31 24	31	0 0.60
0.60	SFC 2 167.00	OOR40 2 209.00	DA 4 42.13	MALE 200.00 0.42133	240 6	IL 6	CA 100.00	31 27	47	0 0.33

				31						
	SSG	95C34	DA	MALE	180	PA	CA	31	31	0
	6	23	16	69.57	40	31	77.50	30		0.50
0.63	72.18	149.00	54.13	0.54125						
	SFC	29W44	DA	FEMALE	168		MD	31	31	0
	6	24	24	100.00	35	42	120.00	58		0.71
0.80	106.60	137.00	73.69	0.73694						
				09						
	SSG	11B34	DA	MALE	141	PA	CA	31	34	0
	9	32	30	93.75	51	52	101.90	18		0.62
0.78	96.46	110.00	66.90	0.66897						
				46						
	SSG	11C34	DA	MALE	144	CA	CA	33	34	0
	7	32	23	71.88	56	35	62.50	9		0.58
0.56	68.78	111.00	57.42	0.57416						
				65						
	SSG	11B34	DA	MALE	180	MS	CA	33	35	0
	9	28	26	92.86	48	51	106.25	6		0.58
0.69	97.28	147.00	61.34	0.61341						
				34						
	SSG	OOR30	DS	MALE	120		MD	34	38	0
	11	26	35	134.62	47	60	127.66	12		0.71
0.73	132.32	86.00	71.66	0.71658						
				3						
	SFC	OOR40	DA	MALE	216		MD	35	38	0
	13	11	8	72.73	16	14	87.50	39		0.57
0.86	77.60	181.00	66.57	0.66571						
				46						
	SGT	OOR20	DS	MALE	108		MD	36	40	0
	15	30	49	163.33	57	83	145.61	35		0.89
0.67	157.49	72.00	81.56	0.81555						
				57						
	SSG	OOR30	DA	FEMALE	168		MD	36	38	0
	9	24	29	120.83	46	59	128.26	13		0.83
0.63	123.28	132.00	76.29	0.76290						
				36						
	SSG	OOR30	DA	MALE	168	OR		CA	36	40
0		10	39	43	110.26	57	71	124.56	63	
0.67	0.56	114.98	132.00	63.00	0.63000					
					03					
	SFC	OOR40	DA	MALE	185		MD	37	36	0
	14	4	7	175.00	15	14	93.33	28		0.75
0.17	148.05	148.00	55.75	0.55750						
				01						
	SGT	13B24	VOL	MALE	156		MD	37	40	0
	13	28	30	107.14	58	65	112.07	1		0.67
0.74	108.77	119.00	69.11	0.69111						
				14						
	SSG	OOR30	DA	MALE	175		MD	39	40	0
	11	31	42	135.48	64	73	114.06	-25		0.80
0.63	128.41	136.00	74.23	0.74225						
	SSG	27G34	DA	MALE	157		MD	40	40	0
	13	43	47	109.30	69	80	115.94	15		0.70
0.50	111.49	117.00	63.40	0.634						
	SSG	OOR30	VOL	MALE	221		MD	41	39	0
	16	18	22	122.22	33	40	121.21	3		0.83
0.75	121.89	180.00	80.58	0.80583						
				31						
	SFC	OOR40	DA	MALE	209		MD	41	39	0
	6	32	36	112.50	60	64	106.67	0		0.64
0.82	110.58	168.00	69.64	0.69636						
				39						
	SFC	OOR30	DA	MALE	203	CA	CA	43	24	0
	7	2	2	100.00	7	2	28.57	16		
76.43	160.00		0							

	SSG	11M34	VOL	MALE	101	SD	CA	44	51	0
	15	41	47	114.63	77	80	103.90	19		0.67
0.67	111.09	57.00	66.67	0.66666						
				67						
	SGT	16S24	DA	MALE	111		MD	45	40	0
	12	35	38	108.57	68	61	89.71	14		0.72
0.42	102.35	66.00	62.35	0.62350						
				42						
	SGT	52D24	DA	MALE	96		MD	45	40	0
	7	40	45	112.50	60	69	115.00	47		0.75
0.67	113.33	51.00	72.25	0.72250						
				01						
	SFC	OOR40	DA	MALE	162	IL	CA	45	43	0
	11	36	40	111.11	60	66	110.00	22		0.62
0.75	110.74	117.00	66.13	0.66132						
				34						
	SFC	OOR40	VOL	MALE	144	WI	CA	45	47	0
	11	30	26	86.67	48	48	100.00	49		0.55
0.68	91.07	99.00	59.54	0.59544						
				47						
	SSG	OOR30	DA	MALE	103		MD	46	40	0
	18	40	33	82.50	66	85	128.79	31		0.66
0.96	97.78	57.00	75.85	0.75850						
				01						
	SGT	62B24	DA	MALE	108		MD	46	36	0
	19	35	37	105.71	58	86	148.28	19		0.63
0.72	119.76	62.00	65.71	0.65708						
				32						
	SFC	OOR40	DA	MALE	96		MD	46	38	0
	14	24	18	75.00	38	44	115.79	39		0.52
0.88	88.46	50.00	63.72	0.63715						
	SSG	OOR30	DA	MALE	204	CA	CA	47	52	0
	25	27	22	81.48	46	32	69.57	35		0.20
0.50	77.55	157.00	29.90	0.299						
	SSG	OOR30	DA	MALE	108		MD	49	40	0
	20	38	53	139.47	64	85	132.81	24		0.77
0.70	137.28	59.00	74.64	0.74642						
				25						
	SSG	OOR30	DA	MALE	130	MD	CA	50	51	0
	19	45	60	133.33	79	111	140.51	-11		0.94
0.64	135.70	80.00	84.18	0.84178						
				79						
	SSG	OOR30	DA	MALE	120	CA	CA	51	50	0
	20	41	45	109.76	67	79	117.91	-11		0.69
0.64	112.45	69.00	67.06	0.67062						
				51						
	SFC	OOR40	VOL	MALE	108		MD	51	39	0
	12	35	54	154.29	53	106	200.00	32		0.82
0.67	169.37	57.00	76.82	0.76818						
				20						
	SFC	OOR40	DA	MALE	144	NY	CA	51	49	0
	21	35	52	148.57	57	107	187.72	5		0.76
0.72	161.49	93.00	74.52	0.74517						
				59						
	SFC	OOR40	VOL	MALE	124		MD	52	40	0
	15	32	42	131.25	63	79	125.40	25		0.69
0.64	129.32	72.00	67.65	0.67647						
				74						
	SSG	19K34	DA	MALE	190		MD	53	40	0
	12	31	22	70.97	57	66	115.79	60		0.66
0.80	85.76	137.00	70.43	0.70428						
				57						
	SSG	31G34	DA	FEMALE	171		MD	53	38	0
12	20	24	120.00	35	52	148.57	19		0.79	0.63
129.43	118.00	73.88	0.7388379							

	SSG	OOR30	DA	MALE	117		MD	53	40	0
	12	34	44	129.41	60	60	100.00	12		0.68
0.57	119.71	64.00	63.98	0.63975						
				71						
	SSG	OOR30	VOL	MALE	132	TX	CA	54	47	0
	13	25	21	84.00	43	42	97.67	15		0.55
0.64	88.51	78.00	57.76	0.57759						
				76						
	SFC	OOR40	DS	MALE	156		MD	54	40	0
	10	32	44	137.50	59	80	135.59	16		0.62
0.78	136.87	102.00	67.05	0.67049						
				02						
	SSG	OOR30	DA	MALE	201		MD	55	40	0
	16	41	44	107.32	62	64	103.23	40		0.80
0.59	105.97	146.00	73.01	0.73011						
				75						
	SFC	OOR40	DA	MALE	131	LA	CA	56	48	0
	15	28	36	128.57	45	52	115.56	14		0.76
0.89	124.28	75.00	80.38	0.80380						
				97						
	SFC	OOR40	DA	MALE	132	CA	CA	57	50	0
	24	36	44	122.22	63	88	139.68	22		0.77
0.86	127.98	75.00	80.16	0.80156						
				71						
	SFC	OOR40	DA	MALE	159		MD	58	40	0
	21	37	48	129.73	60	76	126.67	16		0.83
0.65	128.72	101.00	77.36	0.77355						
				05						
	SFC	OOR40	DA	MALE	169	CA	CA	60	50	0
	14	33	30	90.91	56	47	83.93	27		0.56
0.83	88.61	109.00	64.72	0.64722						
				24						
	SFC	OOR40	DA	MALE	180	KY	CA	61	47	0
	15	49	44	89.80	84	87	103.57	13		0.61
0.76	94.34	119.00	65.98	0.65978						
				93						
	SFC	OOR40	VOL	MALE	192	SC	CA	61	50	0
	24	43	45	104.65	75	100	133.33	31		0.63
0.83	114.12	131.00	69.38	0.69375						
				12						
	SFC	OOR40	DS	MALE	147		MD	63	40	0
	20	28	27	96.43	53	68	128.30	39		0.88
0.75	106.95	84.00	83.63	0.83628						
				79						
	SFC	OOR40	DA	MALE	210		MD	65	38	0
	14	21	25	119.05	36	43	119.44	20		
	SFC	OOR40	DA	MALE	168		MD	65	37	0
	21	14	25	178.57	28	53	189.29	51		0.69
0.84	182.11	103.00	74.00	0.73996						
				35						
	SFC	OOR40	VOL	MALE	219		MD	66	40	0
	12	28	18	64.29	56	78	139.29	-3		0.70
0.90	89.04	153.00	76.50	0.76503						
				44						
	SFC	OOR40	DA	MALE	184	PA	CA	66	46	0
	17	24	26	108.33	40	40	100.00	11		0.71
0.50	105.58	118.00	64.36	0.64357						
				16						
	SFC	OOR40	DA	MALE	48		MD	67	38	0
	15	8	4	50.00	16	17	106.25	43		0.73
0.67	68.56	-19.00	70.73	0.70727						
				30						
	SFC	OOR40	DA	MALE	210	CA	CA	67	45	0
	20	7	24	342.86	17	29	170.59	34		0.78
0.00	286.01	143.00	52.11	0.52111						

				12						
	SFC	OOR40	VOL	MALE	168	PR	CA	68	50	0
	14	19	11	57.89	32	31	96.88	30		0.60
0.83	70.76	100.00	67.70	0.67699						
				98						
	SFC	OOR40	VOL	MALE	192	AL	CA	69	48	0
	33	38	63	165.79	64	97	151.56	19		0.88
0.76	161.09	123.00	83.77	0.83767						
				86						
	SFC	OOR40	DA	MALE	180	VA	CA	74	51	0
	22	53	58	109.43	81	106	130.86	-7		0.76
0.78	116.51	106.00	77.06	0.77061						
				39						
	SFC	00R40	DA	MALE	207	AR	CA	79	6	0
	2	2	4	200.00	7	6	85.71	16		
	162.29	128.00	0.00	0						
	SSG	OOR30	DA	MALE	171		MD	87	19	0
	5	15	16	106.67	27	23	85.19	-28		0.82
0.56	99.58	84.00	73.15	0.73151						
				54						
	SFC	OOR40	VOL	MALE	192		MD	89	40	0
	13	30	21	70.00	53	44	83.02	18		0.53
0.58	74.30	103.00	54.72	0.54720						
				57						
	SSG	OOR30	DA	MALE	192		MD	101	40	0
	14	32	29	90.63	60	56	93.33	5.5		0.61
0.56	91.52	91.00	59.28	0.59277						
				78						
	SFC	OOR40	VOL	MALE	257		MD	102	40	0
	18	31	38	122.58	78	130	166.67	27		0.74
0.81	137.13	155.00	75.85	0.75848						
				04						
	SFC	OOR40	DA	MALE	180		MD	105	38	0
	17	4	6	150.00	6	10	166.67	42		1.00
0.75	155.50	75.00	91.75	0.9175						
	SFC	OOR40	VOL	MALE	238	TX	CA	118	45	0
	7	11	6	54.55	21	17	80.95	26		0.67
0.56	63.26	120.00	63.00	0.63000						
				03						
	SFC	OOR40	VOL	MALE	240		MD	130	38	0
	23	29	39	134.48	45	64	142.22	7		0.78
0.75	137.04	110.00	76.86	0.76861						
				12						
	SSG	71L34	DA	FEMALE	48	NC	CA		12	0
	2	8	11	137.50	16	23	143.75	30		1.00
0.50	139.56	48.00	83.50	0.835						
	SSG	11B34	DA	MALE	118	MO	CA		11	0
	2	6	14	233.33	12	31	258.33	29		0.86
0.67	241.58	118.00	79.43	0.79428						
				59						
	SSG	E35	DA	MALE	160	MO	CA		12	0
	2	9	10	111.11	13	15	115.38	37		0.80
1.00	112.52	160.00	86.60	0.866						

T-TEST

APPENDIX B. HYPOTHESIS TESTS ON UNIT SUCCESS RATES

Two separate units were used in this study's database. Included in this appendix are the hypothesis tests which were performed to determine if the success rates of the Santa Anna Recruiting Battalion and the Baltimore Recruiting Battalion had significantly different means and variances. The tests involved data only for the reduced database which consisted of 101 recruiters.

F-Test Two-Sample for Variances		
Null Hypothesis: Variances are the same for both units		
	Baltimore	Santa Anna
Mean	0.701544314	0.672940284
Variance	0.012786889	0.012865515
Observations	42	59
df	41	58
F	1.006148965	
P(F<=f) one-tail	0.498178449	> 0.10 Accept null
F Critical one-tail	0.69484507	

t-Test: Two-Sample Assuming Equal Variances		
Null Hypothesis: Units have the same mean success		
	Baltimore	Santa Anna
Mean	0.701544314	0.672940284
Variance	0.012786889	0.012865515
Observations	42	59
Pooled Variance	0.012832953	
Hypothesized Mean Difference	0	
df	99	
t Stat	1.250702158	
P(T<=t) one-tail	0.10699552	
t Critical one-tail	1.290161435	
P(T<=t) two-tail	0.213991041	> 0.10 Accept null hypothesis
t Critical two-tail	1.660391717	one-tailed or two-tailed

APPENDIX C. HYPOTHESIS TESTS FOR RECRUITER TIME GROUPS

This appendix includes hypothesis tests for the success rates for the recruiter time groups. The groups were divided into the three different time phases that recruiters are chronologically categorized. A t-test with the TTE recruiters could not be conducted because there were not enough recruiters in this category. The degrees of freedom made comparison impossible. However, the other categories were compared and the results shown.

- TTE Phase: Recruiters with 0-9 months in recruiting duty
- Detailers: Recruiters with 10-24 months of recruiting duty and are not career recruiters (have not changed PMOS to 00R)
- 00R: Recruiters who have over 24 months of recruiting duty and have redesignated their PMOS to become career recruiters

t-Test: Two-Sample Assuming Unequal Variances

Null Hypothesis: Equal means	0-9 Mons	10-24 Mons
Mean	0.59222196	0.67760221
Variance	#DIV/0!	0.01411023
Observations	1	55
Hypothesized Mean Difference	0	
df	0	
t Stat	-5.330546237	
P(T<=t) one-tail	#NUM!	
t Critical one-tail	#NUM!	
P(T<=t) two-tail	#NUM!	
t Critical two-tail	#NUM!	

Unable to perform test because of insufficient data for 0-9 months recruiters

t-Test: Two-Sample Assuming Unequal Variances
Null Hypothesis: Equal Means

	10-24 Mons	24 < Mons
Mean	0.677602214	0.69573320
Variance	0.014110232	0.01163147
Observations	55	45
Hypothesized Mean Difference	0	
df	97	
t Stat	-0.79892611	
P(T<=t) one-tail	0.213142577	
t Critical one-tail	1.290341061	
P(T<=t) two-tail	0.426285155	> 0.10 Accept Null hypothesis
t Critical two-tail	1.66071458	

APPENDIX D. REDUCED DATABASE FOR 101 RECRUITERS

The initial database was reduced from 276 recruiters to 101 recruiters to enable equal comparison of each recruiter's success rate, discussed in Chapter II. The data from the initial database was condensed to represent the information in quantified terms. There are six potential explanatory variables listed, along with the defined success MOE. The column headings are described below.

- AFQT: Score on the Armed Forces Qualification Test
- TEST: Score on the Sales Comprehension Test
- GENDER: Binary variable; Male = 1, Female = 0
- PMOS: Primary Military Occupational Skill; Binary variable; Combat = 1, Noncombat = 0
- SELECTION: Method of assinement to recruiting duty; Binary variable; Volunteer = 1, DA Selected = 0

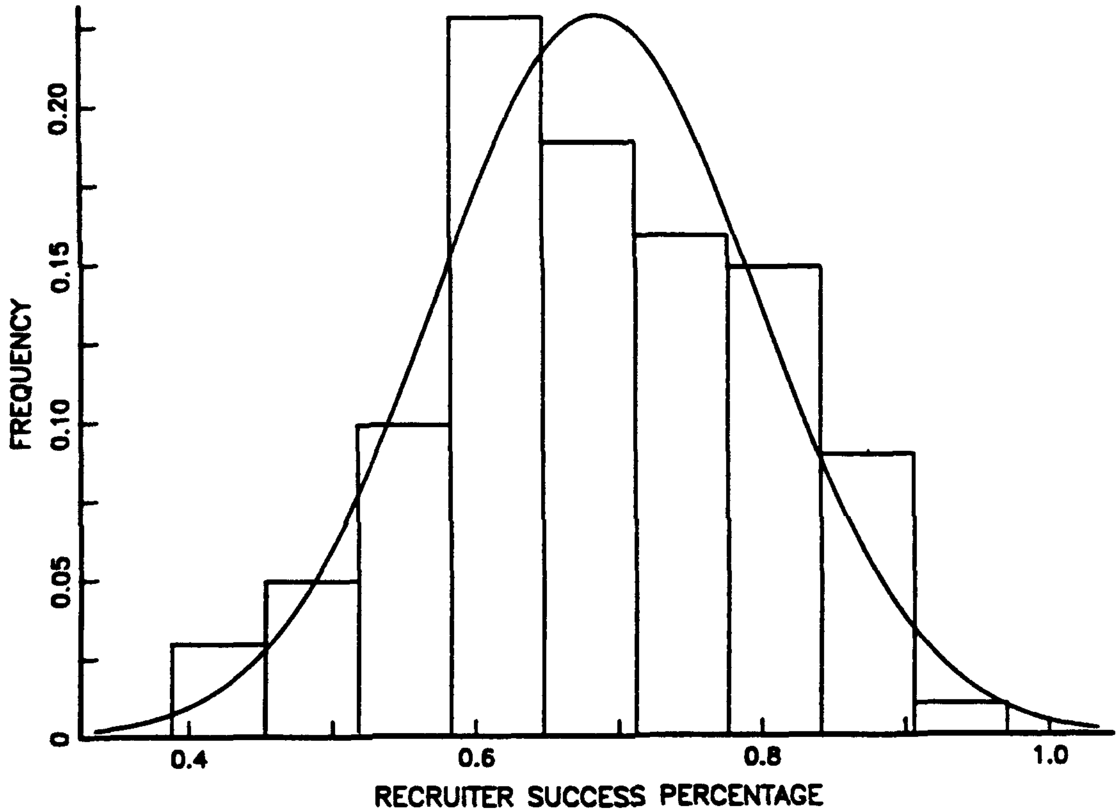
MONTHS: Number of months soldier was in military prior to being assigned to recruiting

AFQT	TEST	GENDER	PMOS	SELECT	MONTHS	SUCCESS RATE
61	4	1	1	0	92	0.7778
54	35	1	1	0	80	0.8511
96	34	1	0	0	61	0.7411
52	20	1	1	0	132	0.5453
95	42	1	0	1	133	0.8522
52	-7	1	1	1	96	0.5911
59	36	1	0	0	130	0.5922
49	31	1	1	0	116	0.8522
43	45	1	1	0	139	0.7274
52	0	1	1	0	127	0.5911
66	42	1	0	0	147	0.5806
36	14	1	0	0	141	0.6678
56	18	1	1	0	103	0.7422
17	32	1	1	0	112	0.7778
99	48	1	0	0	122	0.8900
81	30	1	0	0	85	0.7789
33	28	1	0	1	109	0.6289
40	19	1	0	0	72	0.7033

70	14	1	0	0	68	0.6942
96	36	1	1	0	74	0.7295
50	27	1	0	0	76	0.7044
86	0	1	0	0	113	0.9256
42	49	1	0	0	28	0.7400
17	-1	1	0	1	94	0.8763
38	26	1	1	0	125	0.7789
85	48	1	1	0	72	0.7033
26	9	1	1	0	110	0.4811
72	31	1	0	0	141	0.7686
58	20	1	1	0	115	0.5933
66	24	1	0	0	102	0.6300
32	-23	1	0	0	98	0.7033
77	21	1	1	0	123	0.7400
46	-28	1	0	0	152	0.7044
32	12	1	1	0	112	0.5567
25	20	1	1	0	163	0.6117
85	27	1	0	0	112	0.8350
87	18	1	0	0	80	0.7606
36	31	0	0	0	136	0.6289
38	26	1	1	0	110	0.7411
44	8	1	1	0	93	0.6300
93	35	1	1	0	155	0.8144
23	27	1	1	0	142	0.4433
66	31	1	0	0	125	0.6117
68	40	1	0	0	101	0.7789
25	3	1	0	0	127	0.6117
52	-15	1	1	0	82	0.6197
80	10	1	0	1	74	0.7789
98	28	1	1	0	81	0.5922
76	47	1	0	0	133	0.7411
74	26	1	0	0	64	0.8889
89	44	1	0	0	136	0.7606
75	25	1	0	0	130	0.6678
79	39	1	0	0	109	0.6079
40	13	1	0	0	128	0.5956
82	14	1	1	1	70	0.5911
85	17	0	0	0	86	0.5911
35	33	1	1	0	68	0.7022
78	-3	1	1	0	106	0.6667
93	23	1	0	0	68	0.5922
24	26	1	0	0	83	0.7033
85	6	1	0	0	58	0.5137
79	32	1	1	0	108	0.8156
83	-6	1	0	0	152	0.7789
33	28	1	0	0	103	0.6667
66	0	1	0	0	100	0.5189
23	37	1	0	0	138	0.5704
78	24	1	1	1	96	0.7767
82	15	1	0	0	105	0.8156
50	21	1	1	0	88	0.5785

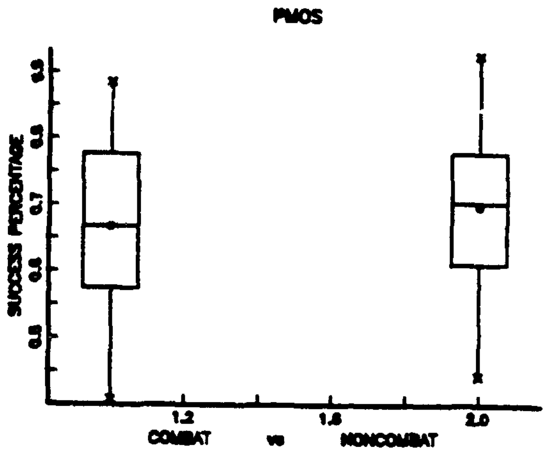
61	9	1	0	0	94	0.6300
44	17	1	0	0	88	0.8522
97	50	1	1	0	69	0.5922
82	31	1	1	0	97	0.5167
75	24	1	1	0	53	0.4078
72	-18	1	1	0	102	0.6289
68	36	1	0	0	116	0.8040
72	-3	1	0	0	80	0.5556
86	56	1	1	0	104	0.7606
25	-12	1	0	0	116	0.7033
71	47	1	0	0	94	0.6667
41	53	1	0	0	137	0.8144
84	30	1	0	0	149	0.5911
32	3	1	0	0	125	0.8522
49	58	0	0	0	137	0.7107
85	18	1	1	0	107	0.7778
90	38	1	1	0	49	0.7411
30	5	0	0	0	142	0.4433
60	29	1	1	0	104	0.5785
52	15	1	1	0	94	0.5178
36	24	1	1	0	151	0.7044
63	20	1	1	1	130	0.7767
41	24	1	1	0	125	0.8843
59	26	1	0	0	97	0.5167
32	20	1	0	0	77	0.7033
17	49	1	0	0	137	0.5189
16	6	1	1	0	145	0.5556
49	23	1	0	0	115	0.7606
61	24	1	0	0	43	0.6300
78	20	1	1	0	126	0.6678
20	57	1	1	1	118	0.6678
32	17	1	0	0	138	0.8156

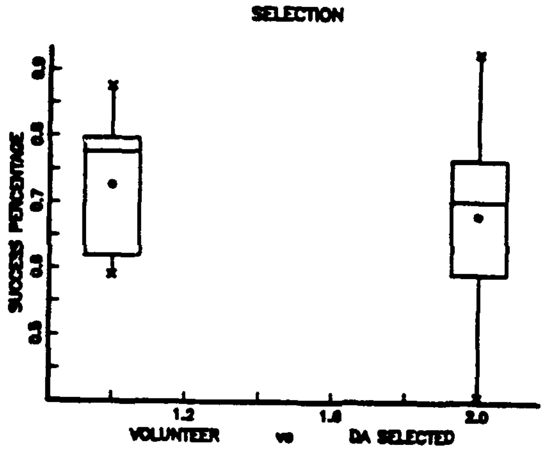
APPENDIX E. FREQUENCY HISTOGRAM FOR RECRUITER SUCCESS RATES



The monthly success MOE discussed in Chapter II possesses a frequency histogram that has a distinctive normal pattern. As the graph shows, the distribution of the success rates for 101 recruiters are centered about a mean of approximately 0.70 and tapers off on either side. Although the graph doesn't have perfect symmetry, the resulting figure is bell-shaped, indicating a normal distribution for the MOE.

APPENDIX F. SUBGROUP BOXPLOTS OF NONINTELLECTIVE GENERAL FACTORS





The box plots shown in each of the graphs reveal that the subgroups of each nonintellective general factor have different means. The difference in means of each categorical factor is evident in the different locations of the center in each subgroup's box.

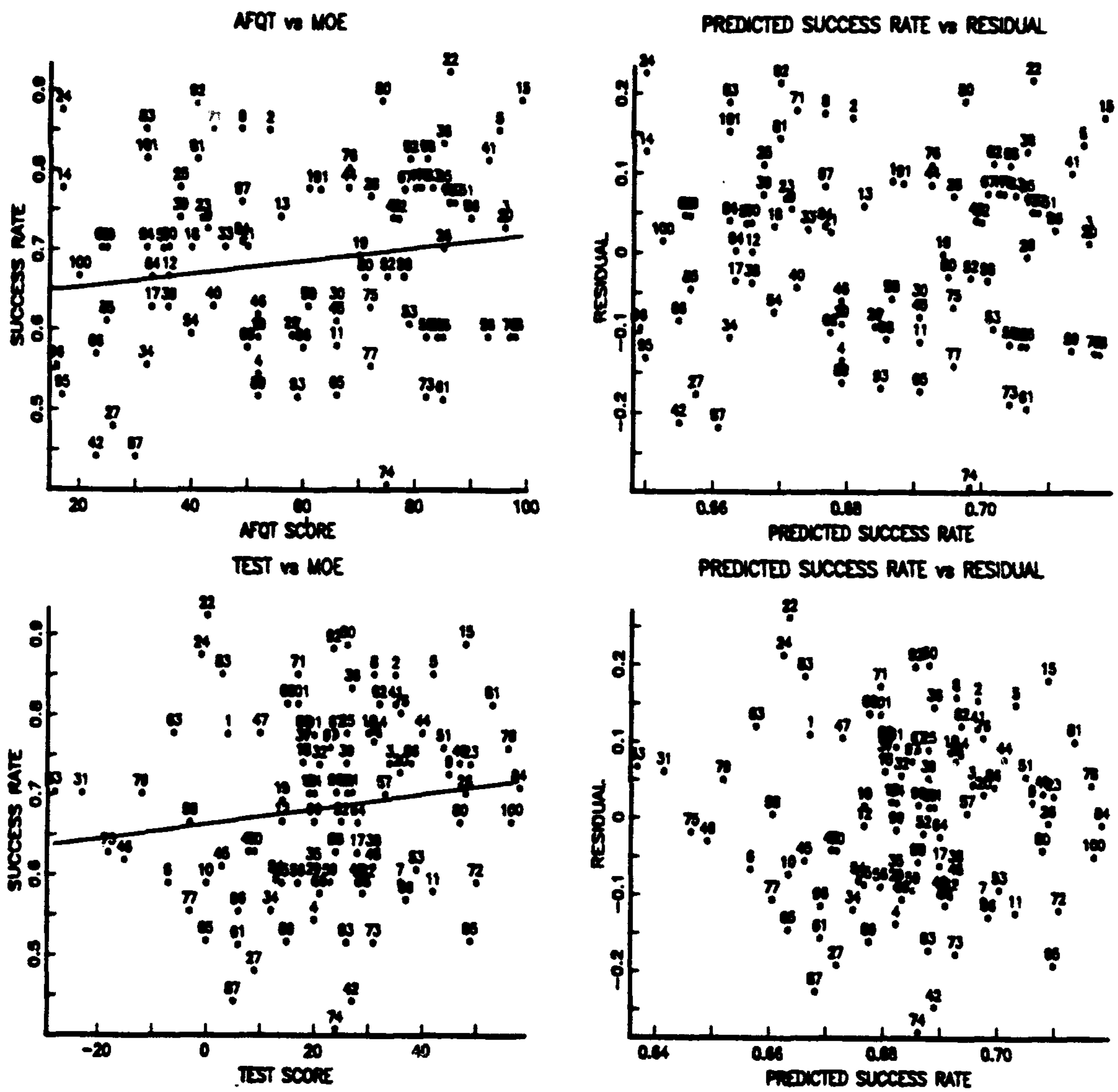
APPENDIX G. EXPECTANCY CHARTS FOR NONINTELLECTIVE GENERAL FACTORS

Three non-intellective general factors described in Chapter II are to be considered as potential predictive variables. The expectancy charts shown below all support this consideration.

PMOS SUBGROUPS	# ABOVE SAMPLE AVG	# BELOW SAMPLE AVG	% ABOVE SAMPLE AVG	SUBGROUP AVG SUCCESS
NCBT	33	24	0.5789	0.6987
CBT	20	24	0.4545	0.6669
GENDER SUBGROUPS	# ABOVE SAMPLE AVG	# BELOW SAMPLE AVG	% ABOVE SAMPLE AVG	SUBGROUP AVG SUCCESS
FEM	1	3	0.2500	0.5935
MALE	52	45	0.5361	0.6886
SELECTION SUBGROUPS	# ABOVE SAMPLE AVG	# BELOW SAMPLE AVG	% ABOVE SAMPLE AVG	SUBGROUP AVG SUCCESS
DA	48	44	0.5217	0.6807
VOL	5	4	0.5556	0.7266

As the charts show, each categorical factor's expected pattern coincides with the subgroup average success pattern, thus supporting the consideration of each factor as a potential predictive variable.

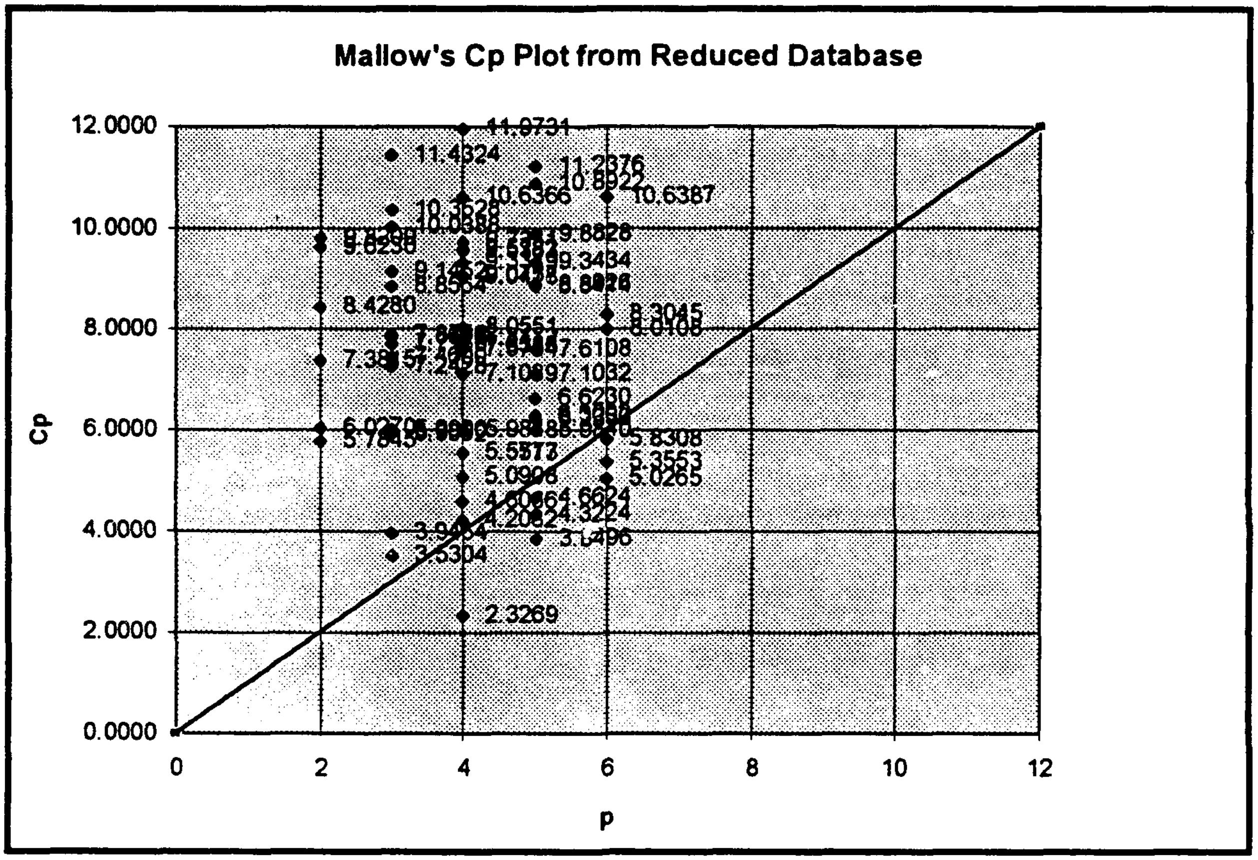
APPENDIX H. GRAPHICAL ANALYSIS FOR OUTLIERS



Two scatterplots and two residual plots from simple linear regression equations are shown in the above graphs. A scatterplot between success rates and AFQT scores with a fitted linear equation in the first graph shows the dispersion of the actual data points. A similar plot with success rates and SCT scores is also shown on the lower left. Residual plots from the simple linear regression models for the corresponding independent variable are shown beside each scatter plot. The data points have been labeled to distinguish unusually

APPENDIX I. MODEL DEVELOPMENT WITHOUT POSSIBLE OUTLIERS

The reduced database has 99 data points. The same method that was used to develop and refine the model from the full database will be employed. Mallow's coefficient was computed for each subset, resulting in the plot shown below.



There are seven subsets which may be considered as the best description of the success MOE. The corresponding variable subsets with the data points below the 45 degree line are shown with their statistics.

VARIABLES	Cp	R	R^2
ATGPS	5.0265	0.3457	0.1195
ATGPM	5.3553	0.3481	0.1211
ATGSM	5.8308	0.3308	0.1094
ATGP	4.6624	0.3418	0.1168
ATGS	3.8496	0.3242	0.1051
ATGM	4.3224	0.3283	0.1078
ATG	2.3269	0.3218	0.1035

The subset with the greatest R is ATGPM. The ANOVA table below shows that the subset does explain some of the variabilty in the MOE. The F-statistic is large enough to reject the null hypothesis that the coefficients of all the variables in the model are equal to zero.

	df	SS	MS	F	p-value
Regression	5	0.1426	0.0285	2.5640	0.0321
Residual	93	1.0345	0.0111		
Total	98	1.1771			

The t-statistics of the model's coefficients are shown below.

	Coefficients	Standard Error	t Stat	**p-value**	Lower 90%	Upper 90%
Intercept	0.48372284	0.0814	5.9390	0.0000	0.3484	0.6190
AFQT	0.00097831	0.0005	2.0205	0.0462	0.0002	0.0018
TEST	0.00097078	0.0006	1.5848	**0.1164**	0.0000	0.0020
GENDER	0.10859848	0.0554	1.9597	0.0530	0.0165	0.2007
PMOS	-0.0259085	0.0218	-1.1902	**0.2370**	-0.0621	0.0103
MONTHS	0.00027084	0.0004	0.6775	**0.4998**	-0.0004	0.0009

It can be seen that the variable Month is the least significant, followed by PMOS and Test. Deletion of the variable Month results in a subset from Mallow's method (**ATGP**) with R equal to 0.3418. The ANOVA table for this model is

	df	SS	MS	F	p-value
Regression	4	0.1375	0.0344	3.1081	0.0189
Residual	94	1.0396	0.0111		
Total	98	1.1771			

As the p-value indicates, this model also explains variabilty of the MOE. The related t-statistics of the coefficients in the model are listed in the following table.

	Coefficients	Standard Error	t Stat	p-value	Lower 90%	Upper 90%
Intercept	0.52180833	0.0588	8.8789	0.0000	0.4242	0.6194
AFQT	0.00088461	0.0005	1.9119	0.0589	0.0001	0.0017
TEST	0.0009896	0.0006	1.6218	**0.1082**	0.0000	0.0020
GENDER	0.10446135	0.0549	1.9020	0.0602	0.0132	0.1957
PMOS	-0.0257882	0.0217	-1.1881	**0.2378**	-0.0618	0.0103

The least significant variable is PMOS which has a p-value of nearly 0.24. Elimination of this variable again results in a subset from Mallow's method. The model **ATG** has a correlation coefficient of 0.3218. Its corresponding ANOVA table and t-statistics are shown below.

	df	SS	MS	F	p-value
Regression	3	0.1219	0.0406	3.6577	0.0152
Residual	95	1.0552	0.0111		
Total	98	1.1771			

	Coefficients	Standard Error	t Stat	**p-value**	Lower 90%	Upper 90%
Intercept	0.5205	0.0589	8.8393	0.0000	0.4227	0.6183
AFQT	0.0009126	0.0005	1.9707	0.0517	0.0001	0.0017
TEST	0.0009857	0.0006	1.6120	**0.1103**	0.0000	0.0020
GENDER	0.09250	0.0541	1.7096	0.0906	0.0026	0.1824

The ANOVA table indicates that the model indeed explains variability in the MOE. All the coefficients of the model are significant with the exception of the variable Test, but all are significant when a one-tailed test is applied to Test. Further, deletion of this variable results in a subset **not** identified by Mallow's method to describe the MOE best. The Test variable's relationship with the other remaining remaining variables increases the amount of variability that the model can describe. As a result the Test variable is retained and the final model is

$$MOE = \beta_0 + \beta_A X_A + \beta_T X_T + \beta_G X_G, \qquad \textbf{(Equation 26)}$$

where,
$\beta_0 = 0.5205$
$\beta_A = 0.0009126$
$\beta_T = 0.0009857$
$\beta_G = 0.09250$
and **r = 0.3218**

This model has three variables as opposed to the four-variable model calculated using the full database. The effects of the data points which might have been outliers are notable. The **ATG** model has a greater correlation coefficient and a smaller MSE. All but one of the coefficients in the ATG model are significant, while three of the four coeffiecients in the **ATGP** model are insignificant. Since the data points in question were high leverage points for AFQT and Test, these variables are most affected by their inclusion or exclusion. Without the " outliers " AFQT becomes significant. The probability that the Test coefficient is equal to zero decreases from over 0.17 down to almost 0.11. Under a one-tailed test both AFQT and Test are completely significant. This model from the reduced data set is an alternative to the model from the full database. The user must beware of deleting data. Although the correlation may be more desirable, the chance of using a model that may not truly describe the system is possible. The study of this model indicates that further investigation of the raw data is warranted, and the study sponsor may desire to further investigate the records of the two recruiters in question.

APPENDIX J. REGRESSION CALCULATIONS FOR MALLOW'S SUBSETS

This appendix includes the regression calculations for each combination of variables which can result from the six independent variables being considered for the model. These calculations serve a two-fold purpose. The F-statistic of the regression model is used to calculate the C_p for the complement of the regression model while providing information on the suitability and form of the candidate model.

Each regression model includes the multiple correlation coefficient, the adjusted and unadjusted coefficient of determination, sum of squares, mean sum of squares, and the statistics of the estimated coefficients.

SUMMARY

Regression Statistics	
Multiple R	0.173260
R Square	0.030019
Adjusted R Square	0.020221
Standard Error	0.112447
Observation	101

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	0.038740	0.038740	3.063859	0.083149
Residual	99	1.251795	0.012644		
Total	100	1.290536			

	Coefficients	*Standard Error*	*t Stat*	*p-value*	*Lower 95%*	*Upper 95%*	*Lower 90.000%*	*Upper 90.000%*
Intercept	0.635758	0.030187	21.06023	2.37494E	0.575859	0.695657	0.585635	0.685881
AFQT	0.000834	0.000476	1.750388	0.083149	-0.000111	0.001779	4.28826E	0.001625

SUMMARY

Regression Statistics	
Multiple R	0.147437
R Square	0.021737
Adjusted R Square	0.011856
Standard Error	0.112926

Observation	101

ANOVA

	df	SS	MS	F	Significance F
Regression	1	0.028053	0.028053	2.199870	0.141198
Residual	99	1.262482	0.012752		
Total	100	1.290536			

	Coefficients	Standard Error	t Stat	p-value	Lower 95%	Upper 95%	Lower 90.000%	Upper 90.000%
Intercept	0.663436	0.018286	36.27972	5.45299E	0.627151	0.699721	0.633073	0.693799
SCT	0.000947	0.000638	1.483195	0.141198	-0.000320	0.002215	-0.000113	0.002008

SUMMARY

Regression Statistics	
Multiple R	0.164080
R Square	0.026922
Adjusted R Square	0.017093
Standard Error	0.112626
Observation	101

ANOVA

	df	SS	MS	F	Significance F
Regression	1	0.034744	0.034744	2.739058	0.101089
Residual	99	1.255792	0.012684		
Total	100	1.290536			

	Coefficients	Standard Error	t Stat	p-value	Lower 95%	Upper 95%	Lower 90.000%	Upper 90.000%
Intercept	0.593500	0.056313	10.53924	7.31125E	0.481762	0.705237	0.499997	0.687002
GDR	0.095101	0.057462	1.655010	0.101089	-0.018917	0.209119	-0.000309	0.190511

SUMMARY

Regression Statistics	
Multiple R	0.129472
R Square	0.016763
Adjusted R Square	0.006831
Standard Error	0.113213
Observation	101

ANOVA

	df	SS	MS	F	Significance F
Regression	1	0.021633	0.021633	1.687849	0.196902
Residual	99	1.268902	0.012817		
Total	100	1.290536			

	Coefficients	Standard Error	t Stat	p-value	Lower 95%	Upper 95%	Lower 90.000%	Upper 90.000%
Intercept	0.697693	0.014995	46.52705	4.28857E	0.667939	0.727447	0.672795	0.722591
PMOS	-0.029516	0.022719	-1.299172	0.196902	-0.074596	0.015563	-0.067239	0.008206

SUMMARY

Regression Statistics	
Multiple R	0.115618
R Square	0.013367
Adjusted R Square	0.003401
Standard Error	0.113408
Observation	101

ANOVA

	df	SS	MS	F	Significance F
Regression	1	0.017251	0.017251	1.341328	0.249586
Residual	99	1.273284	0.012861		
Total	100	1.290536			

	Coefficients	Standard Error	t Stat	p-value	Lower 95%	Upper 95%	Lower 90.000%	Upper 90.000%
Intercept	0.680747	0.011823	57.57509	6.27146E	0.657286	0.704208	0.661115	0.700379
SELECT	0.045873	0.039608	1.158157	0.249586	-0.032719	0.124465	-0.019892	0.111639

SUMMARY

Regression Statistics	
Multiple R	0.023344
R Square	0.000544
Adjusted R Square	-0.009550
Standard Error	0.114142
Observation	101

ANOVA

	df	SS	MS	F	Significance F
Regression	1	0.000703	0.000703	0.053981	0.816754
Residual	99	1.289833	0.013028		

Total	100	1.290536

	Coefficients	Standard Error	t Stat	p-value	Lower 95%	Upper 95%	Lower 90.000%	Upper 90.000%
Intercept	0.674853	0.044437	15.18670	1.31819E	0.586680	0.763026	0.601070	0.748636
MONTHS	9.36149E	0.000402	0.232339	0.816754	-0.000705	0.000893	-0.000575	0.000762

SUMMARY

Regression Statistics	
Multiple R	0.309200
R Square	0.095604
Adjusted R Square	0.048005
Standard Error	0.110841
Observation	101

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	5	0.123381	0.024676	2.008517	0.084352
Residual	95	1.167154	0.012285		
Total	100	1.290536			

	Coefficients	Standard Error	t Stat	p-value	Lower 95%	Upper 95%	Lower 90.000%	Upper 90.000%
Intercept	0.538178	0.076348	7.048991	2.86504E	0.386608	0.689749	0.411360	0.664997
SCT	0.001048	0.000628	1.668685	0.098472	-0.000198	0.002296	4.79811E	0.002092
GDR	0.118248	0.058147	2.033590	0.044781	0.002810	0.233685	0.021662	0.214833
PMOS	-0.039687	0.022656	-1.751683	0.083055	-0.084667	0.005291	-0.077321	-0.002053
SELECT	0.049111	0.038924	1.261709	0.210142	-0.028163	0.126386	-0.015544	0.113767
MONTHS	0.000209	0.000395	0.529729	0.597535	-0.000575	0.000993	-0.000447	0.000865

SUMMARY

Regression Statistics	
Multiple R	0.262862
R Square	0.069096
Adjusted R Square	0.030308
Standard Error	0.111866
Observation	101

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	4	0.089171	0.022292	1.781404	0.138857
Residual	96	1.201364	0.012514		

Total	100	1.290536

	Coefficients	*Standard Error*	*t Stat*	*p-value*	*Lower 95%*	*Upper 95%*	*Lower 90.000%*	*Upper 90.000%*
Intercept	0.567452	0.074993	7.566726	2.31952E	0.418591	0.716312	0.442897	0.692006
GDR	0.112557	0.058584	1.921281	0.057662	-0.003732	0.228846	0.015255	0.209858
PMOS	-0.039234	0.022864	-1.715918	0.089402	-0.084620	0.006152	-0.077210	-0.001258
SELECT	0.047149	0.039266	1.200735	0.232808	-0.030795	0.125093	-0.018068	0.112366
MONTHS	0.000207	0.000398	0.521438	0.603262	-0.000583	0.000999	-0.000454	0.000870

SUMMARY

Regression Statistics	
Multiple R	0.182488
R Square	0.033302
Adjusted R Square	0.003404
Standard Error	0.113408
Observation	101

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	3	0.042977	0.014325	1.113860	0.347346
Residual	97	1.247558	0.012861		
Total	100	1.290536			

	Coefficients	*Standard Error*	*t Stat*	*p-value*	*Lower 95%*	*Upper 95%*	*Lower 90.000%*	*Upper 90.000%*
Intercept	0.682603	0.045696	14.93788	6.81322E	0.591909	0.773298	0.606715	0.758492
PMOS	-0.031594	0.022826	-1.384092	0.169506	-0.076898	0.013710	-0.069502	0.006314
SELECT	0.050552	0.039767	1.271217	0.206691	-0.028374	0.129480	-0.015489	0.116595
MONTHS	0.000107	0.000400	0.268833	0.788629	-0.000687	0.000903	-0.000557	0.000773

SUMMARY

Regression Statistics	
Multiple R	0.119206
R Square	0.014210
Adjusted R Square	-0.005907
Standard Error	0.113936
Observation	101

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	2	0.018338	0.009169	0.706335	0.495943

Residual	98	1.272197	0.012981
Total	100	1.290536	

	Coefficients	Standard Error	t Stat	p-value	Lower 95%	Upper 95%	Lower 90.000%	Upper 90.000%
Intercept	0.668271	0.044714	14.94516	5.12792E	0.579536	0.757006	0.594020	0.742522
SELECT	0.046436	0.039840	1.165544	0.246627	-0.032626	0.125499	-0.019721	0.112593
MONTHS	0.000116	0.000402	0.289405	0.772882	-0.000682	0.000915	-0.000552	0.000785

SUMMARY

Regression Statistics	
Multiple R	0.318953
R Square	0.101731
Adjusted R Square	0.054454
Standard Error	0.110465
Observation	101

ANOVA

	df	SS	MS	F	Significance F
Regression	5	0.131288	0.026257	2.151805	0.065886
Residual	95	1.159248	0.012202		
Total	100	1.290536			

	Coefficients	Standard Error	t Stat	p-value	Lower 95%	Upper 95%	Lower 90.000%	Upper 90.000%
Intercept	0.495481	0.083574	5.928622	4.90717E	0.329565	0.661397	0.356659	0.634302
AFQT	0.000908	0.000488	1.857808	0.066293	-6.23186E	0.001879	9.62142E	0.001720
GDR	0.107667	0.057910	1.859213	0.066092	-0.007298	0.222633	0.011475	0.203859
PMOS	-0.038274	0.022584	-1.694740	0.093400	-0.083109	0.006560	-0.075788	-0.000760
SELECT	0.049320	0.038792	1.271393	0.206692	-0.027692	0.126333	-0.015115	0.113756
MONTHS	0.000419	0.000410	1.024129	0.308374	-0.000394	0.001233	-0.000261	0.001101

SUMMARY

Regression Statistics	
Multiple R	0.262768
R Square	0.069047
Adjusted R Square	0.030257
Standard Error	0.111869
Observation	101

ANOVA

	df	SS	MS	F	Significance F

Regression	4	0.089107	0.022276	1.780036	0.139134
Residual	96	1.201428	0.012514		
Total	100	1.290536			

	Coefficients	Standard Error	t Stat	p-value	Lower 95%	Upper 95%	Lower 90.000%	Uppor 90.000%
Intercept	0.602129	0.061553	9.782265	4.36663E	0.479947	0.724312	0.499897	0.704362
AFQT	0.000949	0.000494	1.919904	0.057838	-3.21960E	0.001931	0.000128	0.001771
PMOS	-0.030937	0.022519	-1.373816	0.172698	-0.075638	0.013763	-0.068340	0.006464
SELECT	0.052668	0.039243	1.342099	0.182729	-0.025229	0.130566	-0.012510	0.117847
MONTHS	0.000333	0.000412	0.809301	0.420342	-0.000485	0.001152	-0.000351	0.001019

SUMMARY

Regression Statistics	
Multiple R	0.225265
R Square	0.050744
Adjusted R Square	0.021386
Standard Error	0.112380
Observation	101

ANOVA

	df	SS	MS	F	Significance F
Regression	3	0.065487	0.021829	1.728446	0.166223
Residual	97	1.225048	0.012629		
Total	100	1.290536			

	Coefficients	Standard Error	t Stat	p-value	Lower 95%	Upper 95%	Lower 90.000%	Upper 90.000%
Intercept	0.587224	0.060865	9.647828	7.72611E	0.466422	0.708026	0.486143	0.688305
AFQT	0.000960	0.000496	1.932167	0.056256	-2.61148E	0.001946	0.000134	0.001785
SELECT	0.048661	0.039313	1.237776	0.218786	-0.029365	0.126687	-0.016627	0.113949
MONTHS	0.000344	0.000414	0.832432	0.407209	-0.000477	0.001167	-0.000343	0.001033

Regression Statistics	
Multiple R	0.189079
R Square	0.035751
Adjusted R Square	0.016072
Standard Error	0.112685
Observation	101

ANOVA

	df	SS	MS	F	Significance F

Regression	2	0.046138	0.023069	1.816760	0.167983
Residual	98	1 244398	0.012697		
Total	100	1.290536			

	Coefficients	*Standard Error*	*t Stat*	*p-value*	*Lower 95%*	*Upper 95%*	*Lower 90.000%*	*Upper 90.000%*
Intercept	0.595636	0.060649	9.820993	2.96102E	0.475279	0.715993	0.494925	0.696347
AFQT	0.000942	0.000498	1.891595	0.061497	-4.62537E	0.001930	0.000115	0.001769
MONTHS	0.000316	0.000414	0.763269	0.447136	-0.000506	0.001140	-0.000372	0.001005

SUMMARY

Regression Statistics	
Multiple R	0.287678
R Square	0.082759
Adjusted R Square	0.034483
Standard Error	0.111625
Observation	101

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	5	0.106803	0.021360	1.714294	0.138741
Residual	95	1.183732	0.012460		
Total	100	1.290536			

	Coefficients	*Standard Error*	*t Stat*	*p-value*	*Lower 95%*	*Upper 95%*	*Lower 90.000%*	*Upper 90.000%*
Intercept	0.594789	0.061727	9.635792	9.91943E	0.472245	0.717332	0.492257	0.697321
AFQT	0.000833	0.000503	1.657439	0.100729	-0.000164	0.001832	-1.81713E	0.001669
SCT	0.000767	0.000643	1.191706	0.236345	-0.000511	0.002045	-0.000302	0.001837
PMOS	-0.031067	0.022470	-1.382555	0.170042	-0.075677	0.013542	-0.068392	0.006258
SELECT	0.053972	0.039173	1.377794	0.171503	-0.023795	0.131740	-0.011096	0.119041
MONTHS	0.000303	0.000412	0.736066	0.463503	-0.000515	0.001122	-0.000381	0.000988

SUMMARY

Regression Statistics	
Multiple R	0.253581
R Square	0.064303
Adjusted R Square	0 025316
Standard Error	0 112154
Observation	101

ANOVA

	df	*SS*	*MS*	*F*	*Significance*

					F
Regression	4	0.082986	0.020746	1.649344	0.168204
Residual	96	1.207550	0.012578		
Total	100	1.290536			

	Coefficients	*Standard Error*	*t Stat*	*p-value*	*Lower 95%*	*Upper 95%*	*Lower 90.000%*	*Upper 90.000%*
Intercept	0.579862	0.061063	9.496075	1.79985E	0.458652	0.701072	0.478443	0.681281
AFQT	0.000844	0.000505	1.671616	0.097856	-0.000158	0.001848	5.42498E	0.001684
SCT	0.000763	0.000647	1.179460	0.241129	-0.000521	0.002047	-0.000311	0.001837
SELECT	0.049941	0.039249	1.272405	0.206302	-0.027968	0.127850	-0.015247	0.115129
MONTHS	0.000314	0.000414	0.759917	0.449166	-0.000507	0.001137	-0.000373	0.001003

SUMMARY

Regression Statistics	
Multiple R	0.220279
R Square	0.048523
Adjusted R Square	0.019096
Standard Error	0.112511
Observation	101

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	3	0.062621	0.020873	1.648930	0.183159
Residual	97	1.227915	0.012658		
Total	100	1.290536			

	Coefficients	*Standard Error*	*t Stat*	*p-value*	*Lower 95%*	*Upper 95%*	*Lower 90.000%*	*Upper 90.000%*
Intercept	0.588709	0.060859	9.673237	6.80882E	0.467919	0.709498	0.487638	0.689779
AFQT	0.000829	0.000506	1.637094	0.104851	-0.000176	0.001835	-1.19734E	0.001671
SCT	0.000740	0.000648	1.141082	0.256644	-0.000547	0.002028	-0.000337	0.001817
MONTHS	0.000286	0.000415	0.690860	0.491303	-0.000537	0.001110	-0.000402	0.000976

SUMMARY

Regression Statistics	
Multiple R	0.209383
R Square	0.043841
Adjusted R Square	0.024328
Standard Error	0.112211
Observation	101

ANOVA

	df	SS	MS	F	Significance F
Regression	2	0.056579	0.028289	2.246735	0.111163
Residual	98	1.233957	0.012591		
Total	100	1.290536			

	Coefficients	Standard Error	t Stat	p-value	Lower 95%	Upper 95%	Lower 90.000%	Upper 90.000%
Intercept	0.624631	0.031541	19.80356	4.95646E	0.562039	0.687224	0.572255	0.677008
AFQT	0.000728	0.000483	1.505151	0.135501	-0.000231	0.001688	-7.51743E	0.001531
SCT	0.000768	0.000645	1.190258	0.236820	-0.000512	0.002050	-0.000303	0.001841

SUMMARY

Regression Statistics	
Multiple R	0.301193
R Square	0.090717
Adjusted R Square	0.042860
Standard Error	0.111140
Observation	101

ANOVA

	df	SS	MS	F	Significance F
Regression	5	0.117073	0.023414	1.895591	0.102281
Residual	95	1.173462	0.012352		
Total	100	1.290536			

	Coefficients	Standard Error	t Stat	p-value	Lower 95%	Upper 95%	Lower 90.000%	Upper 90.000%
Intercept	0.481427	0.084692	5.684431	1.43824E	0.313291	0.649562	0.340749	0.622105
AFQT	0.000799	0.000501	1.594131	0.114228	-0.000196	0.001795	-3.35671E	0.001632
SCT	0.000834	0.000642	1.298695	0.197192	-0.000441	0.002110	-0.000232	0.001901
GDR	0.095514	0.057496	1.661220	0.099965	-0.018630	0.209660	9.75994E	0.191019
SELECT	0.046247	0.038958	1.187118	0.238140	-0.031093	0.123589	-0.018463	0.110959
MONTHS	0.000390	0.000413	0.945663	0.346719	-0.000429	0.001210	-0.000295	0.001076

SUMMARY

Regression Statistics	
Multiple R	0.277900
R Square	0.077228
Adjusted R Square	0.038780
Standard Error	0.111377
Observation	101

ANOVA

	df	SS	MS	F	Significance F
Regression	4	0.099666	0.024916	2.008614	0.099370
Residual	96	1.190869	0.012404		
Total	100	1.290536			

	Coefficients	Standard Error	t Stat	p-value	Lower 95%	Upper 95%	Lower 90.000%	Upper 90.000%
Intercept	0.485578	0.084800	5.726147	1.17351E	0.317251	0.653905	0.344735	0.626421
AFQT	0.000783	0.000502	1.560000	0.122050	-0.000213	0.001781	-5.06935E	0.001618
SCT	0.000816	0.000643	1.268189	0.207797	-0.000461	0.002094	-0.000252	0.001885
GDR	0.099409	0.057525	1.728110	0.087183	-0.014776	0.213596	0.003867	0.194952
MONTHS	0.000367	0.000413	0.889416	0.376002	-0.000453	0.001188	-0.000319	0.001054

SUMMARY

Regression Statistics	
Multiple R	0.263865
R Square	0.069624
Adjusted R Square	0.040850
Standard Error	0.111257
Observation	101

ANOVA

	df	SS	MS	F	Significance F
Regression	3	0.089853	0.029951	2.419677	0.070760
Residual	97	1.200682	0.012378		
Total	100	1.290536			

	Coefficients	Standard Error	t Stat	p-value	Lower 95%	Upper 95%	Lower 90.000%	Upper 90.000%
Intercept	0.537084	0.061880	8.679326	9.47305E	0.414268	0.659900	0.434317	0.639850
AFQT	0.000657	0.000481	1.365919	0.175122	-0.000297	0.001613	-0.000141	0.001457
SCT	0.000847	0.000642	1.320370	0.189818	-0.000426	0.002122	-0.000218	0.001914
GDR	0.093606	0.057092	1.639559	0.104336	-0.019706	0.206918	-0.001207	0.188420

SUMMARY

Regression Statistics	
Multiple R	0.227441
R Square	0.051729
Adjusted R Square	0.032377
Standard Error	0.111747
Observation	101

ANOVA

	df	SS	MS	F	Significance F
Regression	2	0.066759	0.033379	2.673034	0.074075
Residual	98	1.223777	0.012487		
Total		1.290536			

	Coefficients	Standard Error	t Stat	p-value	Lower 95%	Upper 95%	Lower 90.000%	Upper 90.000%
Intercept	0.565362	0.058572	9.652390	6.87858E	0.449127	0.681596	0.468099	0.662624
SCT	0.001013	0.000633	1.601169	0.112558	-0.000242	0.002270	-3.76046E	0.002065
GDR	0.100555	0.057115	1.760552	0.081433	-0.012789	0.2139	0.005711	0.195399

SUMMARY

Regression Statistics	
Multiple R	0.320173
R Square	0.102511
Adjusted R Square	0.055274
Standard Error	0.110417
Observation	101

ANOVA

	df	SS	MS	F	Significance F
Regression	5	0.132294	0.026458	2.170178	0.063820
Residual	95	1.158242	0.012192		
Total	100	1.290536			

	Coefficients	Standard Error	t Stat	p-value	Lower 95%	Upper 95%	Lower 90.000%	Upper 90.000%
Intercept	0.486531	0.084071	5.787121	9.17199E	0.319628	0.653434	0.346884	0.626178
AFQT	0.000761	0.000498	1.528132	0.129802	-0.000227	0.001750	-6.62409E	0.001589
SCT	0.000832	0.000638	1.303979	0.195391	-0.000434	0.002099	-0.000227	0.001892
GDR	0.116376	0.057964	2.007702	0.047515	0.001301	0.231451	0.020093	0.212658
PMOS	-0.036855	0.022529	-1.635895	0.105170	-0.081582	0.007870	-0.074278	0.000566
MONTHS	0.000365	0.00041	0.891740	0.374785	-0.000448	0.001179	-0.000315	0.001046

ATGP

Regression Statistics	
Multiple R	**0.308218**
R Square	**0.094998**
Adjusted R Square	**0.057290**
Standard Error	**0.110299**

Observation	101

ANOVA

	df	SS	MS	F	Significance F
Regression	4	0.122599	0.030649	2.519297	0.046156
Residual	96	1.167937	0.012166		
Total	100	1.290536			

	Coefficients	Standard Error	t Stat	p-value	Lower 90.000%	Upper 90.000%
Intercept	0.537728	0.061349	8.765001	**6.67150E**	0.435834	0.639623
AFQT	0.000636	0.000477	1.331976	**0.186021**	-0.000157	0.001429
SCT	0.000863	0.000636	1.356409	**0.178150**	-0.000193	0.001921
GDR	0.110638	0.057545	1.922633	**0.057490**	0.015062	0.206213
PMOS	-0.036922	0.022505	-1.640600	**0.104153**	-0.074300	0.000456

SUMMARY

Regression Statistics	
Multiple R	0.279773
R Square	0.078273
Adjusted R Square	0.049766
Standard Error	0.110738
Observation	101

ANOVA

	df	SS	MS	F	Significance F
Regression	3	0.101014	0.033671	2.745762	0.047131
Residual	97	1.189521	0.012263		
Total	100	1.290536			

	Coefficients	Standard Error	t Stat	p-value	Lower 95%	Upper 95%	Lower 90.000%	Upper 90.000%
Intercept	0.565070	0.058043	9.735239	5.00187E	0.449869	0.680271	0.468676	0.661464
SCT	0.001024	0.000627	1.632421	0.105833	-0.000221	0.002270	-1.77568E	0.002066
GDR	0.117735	0.057526	2.046641	0.043398	0.003562	0.231908	0.022200	0.213269
PMOS	-0.037749	0.022586	-1.671342	0.097877	-0.082576	0.007078	-0.075258	-0.000240

SUMMARY

Regression Statistics	
Multiple R	0.230112
R Square	0.052951
Adjusted R Square	0.033624
Standard	0.111675

Error	
Observation	101

ANOVA

	df	SS	MS	F	Significance F
Regression	2	0.068336	0.034168	2.739702	0.069540
Residual	98	1.222200	0.012471		
Total	100	1.290536			

	Coefficients	Standard Error	t Stat	p-value	Lower 95%	Upper 95%	Lower 90.000%	Upper 90.000%
Intercept	0.593500	0.055837	10.62900	5.24609E	0.482691	0.704308	0.500778	0.686221
GDR	0.112057	0.057906	1.935138	0.055856	-0.002856	0.226970	0.015900	0.208213
PMOS	-0.037379	0.022776	-1.641186	0.103964	-0.082578	0.007818	-0.075200	0.000441

ATGPS

Regression Statistics	
Multiple R	**0.331722**
R Square	**0.110039**
Adjusted R Square	**0.063199**
Standard Error	**0.109953**
Observation	**101**

ANOVA

	df	SS	MS	F	Significance F
Regression	5	0.142010	0.028402	2.349262	0.046695
Residual	95	1.148526	0.012089		
Total	100	1.290536			

	Coefficients	Standard Error	t Stat	p-value	Lower 95%	Upper 95%	Lower 90.000%	Upper 90.000%
Intercept	0.536756	0.061161	8.776007	6.80987E	0.415334	0.658177	0.435163	0.638349
AFQT	0.000643	0.000476	1.351298	0.179808	-0.000301	0.001588	-0.000147	0.001434
SCT	0.000885	0.000634	1.394804	0.166327	-0.000374	0.002146	-0.000169	0.001940
GDR	0.106999	0.057436	1.862926	0.065562	-0.007025	0.221025	0.011594	0.202404
PMOS	-0.038787	0.022482	-1.725187	0.087745	-0.083420	0.005846	-0.076132	-0.001441
SELECT	0.048888	0.038582	1.267106	0.208214	-0.027708	0.125485	-0.015199	0.112976

SUMMARY

Regression Statistics	
Multiple R	0.304849
R Square	0.092933
Adjusted R Square	0.055139

Standard Error	0.110425
Observation	101

ANOVA

	df	SS	MS	F	Significance F
Regression	4	0.119934	0.029983	2.458919	0.050574
Residual	96	1 170602	0.012193		
Total	100	1.290536			

	Coefficients	*Standard Error*	*t Stat*	*p-value*	*Lower 95%*	*Upper 95%*	*Lower 90.000%*	*Upper 90.000%*
Intercept	0.564417	0.057881	9.751203	5.09213E	0.449523	0.679312	0.468282	0.660552
S[illegible]	0.001048	0.000626	1.673871	0.097411	-0.000194	0.002290	8.13242E	0.002087
GDR	0.114223	0.057432	1.988825	0.049567	0.000220	0.228226	0.018834	0.209612
PMOS	-0.039599	0.022571	-1.754433	0.082547	-0.084403	0.005203	-0.077087	-0.002111
SELECT	0.048262	0.038745	1.245613	0.215938	-0.028647	0.125171	-0.016089	0.112614

SUMMARY

Regression Statistics	
Multiple R	0.257798
R Square	0.066459
Adjusted R Square	0.037587
Standard Error	0.111446
Observation	101

ANOVA

	df	SS	MS	F	Significance F
Regression	3	0.085768	0.028589	2.301851	0.081923
Residual	97	1.204767	0.012420		
Total	100	1.290536			

	Coefficients	*Standard Error*	*t Stat*	*p-value*	*Lower 95%*	*Upper 95%*	*Lower 90.000%*	*Upper 90.000%*
Intercept	0.593500	0.055723	10.65086	5.30331E	0.482905	0.704095	0.500959	0.686040
GDR	0 108562	0.057862	1.876198	0.063634	-0.006279	0.223404	0.012468	0.204656
PMOS	-0.039147	0.022778	-1.718617	0.088874	-0.084355	0.006061	-0.076975	-0.001318
SELECT	0.046306	0.039086	1.184729	0.239019	-0.031268	0.123881	-0.018604	0.111217

SUMMARY

Regression Statistics	
Multiple R	0.180504
R Square	0.032581
Adjusted R	0.012838

Square	
Standard Error	0.112870
Observation	101

ANOVA

	df	SS	MS	F	Significance F
Regression	2	0.042048	0.021024	1.650278	0.197286
Residual	98	1.248488	0.012739		
Total	100	1.290536			

	Coefficients	Standard Error	t Stat	p-value	Lower 95%	Upper 95%	Lower 90.000%	Upper 90.000%
Intercept	0.694181	0.015205	45.65404	7.55684E	0.664007	0.724355	0.668932	0.719430
PMOS	-0.031691	0.022715	-1.395136	0.166128	-0.076769	0.013386	-0.069411	0.006029
SELECT	0.050045	0.039533	1.265873	0.208559	-0.028408	0.128498	-0.015603	0.115693

SUMMARY

Regression Statistics	
Multiple R	0.237139
R Square	0.056235
Adjusted R Square	0.016911
Standard Error	0.112637
Observation	101

ANOVA

	df	SS	MS	F	Significance F
Regression	4	0.072573	0.018143	1.430066	0.229920
Residual	96	1.217962	0.012687		
Total	100	1.290536			

	Coefficients	Standard Error	t Stat	p-value	Lower 95%	Upper 95%	Lower 90.000%	Upper 90.000%
Intercept	0.660829	0.047571	13.89115	1.05605E	0.566399	0.755258	0.581817	0.739840
SCT	0.000973	0.000637	1.527340	0.129964	-0.000291	0.002239	-8.51386E	0.002032
PMOS	-0.031656	0.022671	-1.396317	0.165839	-0.076659	0.013345	-0.069311	0.005998
SELECT	0.052535	0.039518	1.329385	0.186871	-0.025908	0.130978	-0.013100	0.118170
MONTHS	0.000104	0.000398	0.262050	0.793844	-0.000685	0.000894	-0.000556	0.000765

SUMMARY

Regression Statistics	
Multiple R	0.192530
R Square	0.037067
Adjusted R	0.007286

Square	
Standard Error	0.113187
Observation	101

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	3	0.047837	0.015945	1.244668	0.297831
Residual	97	1.242698	0.012811		
Total	100	1.290536			

	Coefficients	*Standard Error*	*t Stat*	*p-value*	*Lower 95%*	*Upper 95%*	*Lower 90.000%*	*Upper 90.000%*
Intercept	0.646504	0.046679	13.84987	1.02846E	0.553858	0.739150	0.568983	0.724025
SCT	0.000972	0.000640	1.517418	0.132413	-0.000299	0.002243	-9.18034E	0.002036
SELECT	0.048406	0.039599	1.222397	0.224519	-0.030188	0.127001	-0.017357	0.114171
MONTHS	0.000113	0.000400	0.282797	0.777935	-0.000680	0.000907	-0.000551	0.000777

SUMMARY

Regression Statistics	
Multiple R	0.149111
R Square	0.022234
Adjusted R Square	0.002279
Standard Error	0.113472
Observation	101

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	2	0.028694	0.014347	1.114254	0.332280
Residual	98	1.261842	0.012875		
Total	100	1.290536			

	Coefficients	*Standard Error*	*t Stat*	*p-value*	*Lower 95%*	*Upper 95%*	*Lower 90.000%*	*Upper 90.000%*
Intercept	0.653933	0.046398	14.09378	2.63865E	0.561856	0.746010	0.576885	0.730980
SCT	0.000946	0.000641	1.474411	0.143576	-0.000327	0.002220	-0.000119	0.002012
MONTHS	8.93492E	0.000400	0.223057	0.823955	-0.000705	0.000884	-0.000575	0.000754

SUMMARY

Regression Statistics	
Multiple R	0.257670
R Square	0.066393
Adjusted R Square	0.027493

Standard Error	0.112029
Observation	101

ANOVA

	df	SS	MS	F	Significance F
Regression	4	0.085683	0.021420	1.706773	0.154792
Residual	96	1.204852	0.012550		
Total	100	1.290536			

	Coefficients	Standard Error	t Stat	p-value	Lower 95%	Upper 95%	Lower 90.000%	Upper 90.000%
Intercept ·	0.539180	0.077164	6.987438	3.68201E	0.386010	0.692350	0.411019	0.667341
SCT	0.001035	0.000635	1.630324	0.106309	-0.000225	0.002296	-1.94088E	0.002090
GDR	0.100493	0.057870	1.736520	0.085679	-0.014378	0.215366	0.004377	0.196610
SELECT	0.044607	0.039255	1.136326	0.258648	-0.033314	0.122529	-0.020591	0.109806
MONTHS	0.000204	0.000399	0.511424	0.610228	-0.000588	0.000997	-0.000459	0.000867

SUMMARY

Regression Statistics	
Multiple R	0.232027
R Square	0.053836
Adjusted R Square	0.024573
Standard Error	0.112197
Observation	101

ANOVA

	df	SS	MS	F	Significance F
Regression	3	0.069478	0.023159	1.839762	0.145030
Residual	97	1.221058	0.012588		
Total	100	1.290536			

	Coefficients	Standard Error	t Stat	p-value	Lower 95%	Upper 95%	Lower 90.000%	Upper 90.000%
Intercept	0.542091	0.077237	7.018498	3.05799E	0.388796	0.695386	0.413821	0.670360
SCT	0.001014	0.000635	1.595118	0.113938	-0.000247	0.002276	-4.17079E	0.002070
GDR	0.104159	0.057867	1.799956	0.074976	-0.010691	0.219010	0.008057	0.200260
MONTHS	0.000185	0.000399	0.464744	0.643155	-0.000607	0.000978	-0.000477	0.000849

SUMMARY

Regression Statistics	
Multiple R	0.170346
R Square	0.029017
Adjusted R	0.009201

Square	
Standard Error	0.113077
Observation	101

ANOVA

	df	SS	MS	F	Significance F
Regression	2	0.037448	0.018724	1.464365	0.236239
Residual	98	1.253087	0.012786		
Total	100	1.290536			

	Coefficients	Standard Error	t Stat	p-value	Lower 95%	Upper 95%	Lower 90.000%	Upper 90.000%
Intercept	0.570298	0.075776	7.526110	2.56495E	0.419923	0.720673	0.444468	0.696128
GDR	0.098693	0.058219	1.695206	0.093211	-0.016840	0.214228	0.002017	0.195370
MONTHS	0.000185	0.000402	0.459877	0.646622	-0.000614	0.000984	-0.000483	0.000854

SUMMARY

Regression Statistics	
Multiple R	0.283637
R Square	0.08045
Adjusted R Square	0.042135
Standard Error	0.111182
Observation	101

ANOVA

	df	SS	MS	F	Significance F
Regression	4	0.103823	0.025955	2.099722	0.086775
Residual	96	1.186712	0.012361		
Total	100	1.290536			

	Coefficients	Standard Error	t Stat	p-value	Lower 95%	Upper 95%	Lower 90.000%	Upper 90.000%
Intercept	0.541416	0.076540	7.073635	2.44922E	0.389485	0.693347	0.414292	0.668540
SCT	0.001024	0.000630	1.626302	0.107163	-0.000226	0.002275	-2.17891E	0.002071
GDR	0.121420	0.058271	2.083697	0.039844	0.005752	0.237089	0.024638	0.218203
PMOS	-0.037799	0.022676	-1.666857	0.098802	-0.082812	0.007214	-0.075463	-0.000135
MONTHS	0.000188	0.000396	0.476684	0.634670	-0.000597	0.000974	-0.000469	0.000846

SUMMARY

Regression Statistics	
Multiple R	0.197133
R Square	0.038861
Adjusted R	0.009135

Square	
Standard Error	0.113081
Observation	101

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	3	0.050152	0.016717	1.307325	0.276480
Residual	97	1.240384	0.012787		
Total	100	1.290536			

	Coefficients	*Standard Error*	*t Stat*	*p-value*	*Lower 95%*	*Upper 95%*	*Lower 90.000%*	*Upper 90.000%*
Intercept	0.667826	0.047466	14.06942	3.69518E	0.573618	0.762034	0.588997	0.746654
SCT	0.000945	0.000639	1.478650	0.142473	-0.000323	0.002215	-0.000116	0.002008
PMOS	-0.029401	0.022697	-1.295394	0.198258	-0.074449	0.015645	-0.067095	0.008291
MONTHS	7.92964E	0.000399	0.198607	0.842985	-0.000713	0.000871	-0.000583	0.000742

SUMMARY

Regression Statistics	
Multiple R	0.196139
R Square	0.038470
Adjusted R Square	0.018847
Standard Error	0.112526
Observation	101

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	2	0.049647	0.024823	1.960481	0.146273
Residual	98	1.240888	0.012662		
Total	100	1.290536			

	Coefficients	*Standard Error*	*t Stat*	*p-value*	*Lower 95%*	*Upper 95%*	*Lower 90.000%*	*Upper 90.000%*
Intercept	0.676298	0.020713	32.65052	1.82498E	0.635193	0.717403	0.641903	0.710693
SCT	0.000946	0.000636	1.487428	0.140112	-0.000316	0.002210	-0.000110	0.002003
PMOS	-0.029489	0.022581	-1.305916	0.194638	-0.074301	0.015322	-0.066986	0.008008

SUMMARY

Regression Statistics	
Multiple R	0.294019
R Square	0.086447
Adjusted R Square	0.048382

Standard Error	0.110819
Observation	101

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	4	0.111563	0.027890	2.271064	0.067138
Residual	96	1.178972	0.012280		
Total	100	1.290536			

	Coefficients	*Standard Error*	*t Stat*	*p-value*	*Lower 95%*	*Upper 95%*	*Lower 90.000%*	*Upper 90.000%*
Intercept	0.499545	0.083780	5.962514	4.12050E	0.333241	0.665849	0.360394	0.638695
AFQT	0.000889	0.000490	1.814511	0.072722	-8.35909E	0.001863	7.53310E	0.001704
GDR	0.111085	0.058033	1.914166	0.058577	-0.004109	0.226280	0.014698	0.207471
PMOS	-0.036408	0.022608	-1.610359	0.110602	-0.081286	0.008469	-0.073958	0.001142
MONTHS	0.000394	0.000410	0.961258	0.338837	-0.000420	0.001210	-0.000287	0.001077

SUMMARY

Regression Statistics	
Multiple R	0.278665
R Square	0.077654
Adjusted R Square	0.049128
Standard Error	0.110776
Observation	101

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	3	0.100215	0.033405	2.722211	0.048536
Residual	97	1.190320	0.012271		
Total	100	1.290536			

	Coefficients	*Standard Error*	*t Stat*	*p-value*	*Lower 95%*	*Upper 95%*	*Lower 90.000%*	*Upper 90.000%*
Intercept	0.555547	0.060185	9.230587	6.15209E	0.436095	0.674998	0.455596	0.655498
AFQT	0.000759	0.000470	1.611796	0.110255	-0.000175	0.001693	-2.30373E	0.001541
GDR	0.104652	0.057623	1.816146	0.072436	-0.009713	0.219019	0.008956	0.200349
PMOS	-0.036461	0.022599	-1.613371	0.109912	-0.081316	0.008392	-0.073993	0.001069

SUMMARY

Regression Statistics	
Multiple R	0.215153
R Square	0.046290
Adjusted R	0.026827

Square	
Standard Error	0.112067
Observation	101

ANOVA

	df	SS	MS	F	Significance F
Regression	2	0.059739	0.029869	2.378345	0.098035
Residual	98	1.230796	0.012559		
Total	100	1.290536			

	Coefficients	*Standard Error*	*t Stat*	*p-value*	*Lower 95%*	*Upper 95%*	*Lower 90.000%*	*Upper 90.000%*
Intercept	0.648828	0.031738	20.44315	4.01225E	0.585844	0.711811	0.596125	0.701531
AFQT	0.000827	0.000474	1.741884	0.084666	-0.000115	0.001769	3.86262E	0.001615
PMOS	-0.029082	0.022490	-1.293071	0.199026	-0.073714	0.015550	-0.066429	0.008264

SUMMARY

Regression Statistics	
Multiple R	0.303008
R Square	0.091814
Adjusted R Square	0.053973
Standard Error	0.110493
Observation	101

ANOVA

	df	SS	MS	F	Significance F
Regression	4	0.118489	0.029622	2.426312	0.053130
Residual	96	1.172046	0.012208		
Total	100	1.290536			

	Coefficients	*Standard Error*	*t Stat*	*p-value*	*Lower 95%*	*Upper 95%*	*Lower 90.000%*	*Upper 90.000%*
Intercept	0.555043	0.060033	9.245593	6.21488E	0.435878	0.674209	0.455335	0.654752
AFQT	0.000769	0.000469	1.637097	0.104884	-0.000163	0.001701	-1.11746E	0.001549
GDR	0.100975	0.057555	1.754422	0.082549	-0.013270	0.215221	0.005383	0.196568
PMOS	-0.038259	0.022590	-1.693643	0.093575	-0.083100	0.006581	-0.075778	-0.000740
SELECT	0 047417	0 038757	1.223433	0.224160	-0.029516	0.124351	-0.016954	0.111789

SUMMARY

Regression Statistics	
Multiple R	0.254319
R Square	0.064678
Adjusted R	0.035750

Square	
Standard Error	0.111552
Observation	101

ANOVA

	df	SS	MS	F	Significance F
Regression	3	0.083469	0.027823	2.235872	0.088917
Residual	97	1.207066	0.012443		
Total	100	1.290536			

	Coefficients	Standard Error	t Stat	p-value	Lower 95%	Upper 95%	Lower 90.000%	Upper 90.000%
Intercept	0.554089	0.060606	9.142462	9.53173E	0.433802	0.674375	0.453439	0.654738
AFQT	0.000788	0.000474	1.662299	0.099681	-0.000152	0.001729	7.51329E	0.001575
GDR	0.083841	0.057202	1.465707	0.145962	-0.029688	0.197371	-0.011154	0.178837
SELECT	0.043148	0.039046	1.105050	0.271871	-0.034348	0.120644	-0.021696	0.107993

SUMMARY

Regression Statistics	
Multiple R	0.230007
R Square	0.052903
Adjusted R Square	0.033574
Standard Error	0.111678
Observation	101

ANOVA

	df	SS	MS	F	Significance F
Regression	2	0.068273	0.034136	2.737064	0.069714
Residual	98	1.222262	0.012472		
Total	100	1.290536			

	Coefficients	Standard Error	t Stat	p-value	Lower 95%	Upper 95%	Lower 90.000%	Upper 90.000%
Intercept	0.554590	0.060672	9.140676	8.86779E	0.434186	0.674993	0.453839	0.655340
AFQT	0.000778	0.000474	1.639621	0.104290	-0.000163	0.001720	-9.93364E	0.001566
GDR	0.087937	0.057146	1.538809	0.127073	-0.025467	0.201341	-0.006957	0.182831

ATGS

Regression Statistics	
Multiple R	0.286631
R Square	0.082157
Adjusted R Square	0.043914

Standard Error	0.111079
Observation	101

ANOVA

	df	SS	MS	F	Significance F
Regression	4	0.106027	0.026506	2.148285	0.080706
Residual	96	1.184508	0.012338		
Total	100	1.290536			

	Coefficients	Standard Error	t Stat	p-value	Lower 90.000%	Upper 90.000%
Intercept	0.536168	0.061787	8.677676	**1.02585E**	0.433547	0.638790
AFQT	0.000665	0.000480	1.383777	**0.169635**	-0.000133	0.001463
SCT	0.000867	0.000641	1.352128	**0.179511**	-0.000198	0.001932
GDR	0.089508	0.057113	1.567207	**0.120356**	-0.005350	0.184367
SELECT	0.044531	0.038894	1.144922	**0.255087**	-0.020067	0.109129

SUMMARY

Regression Statistics	
Multiple R	0.242229
R Square	0.058675
Adjusted R Square	0.029561
Standard Error	0.111910
Observation	101

ANOVA

	df	SS	MS	F	Significance F
Regression	3	0.075722	0.025240	2.015413	0.116820
Residual	97	1.214814	0.012523		
Total	100	1.290536			

	Coefficients	Standard Error	t Stat	p-value	Lower 95%	Upper 95%	Lower 90.000%	Upper 90.000%
Intercept	0.619476	0.031731	19.52222	2.27286E	0.556497	0.682455	0.566778	0.672174
AFQT	0.000733	0.000482	1.519320	0.131934	-0.000224	0.001690	-6.82178E	0.001534
SCT	0.000793	0.000644	1.231257	0.221203	-0.000485	0.002072	-0.000276	0.001863
SELECT	0.048350	0.039108	1.236338	0.219317	-0.029267	0.125969	-0.016596	0.113298

SUMMARY

Regression Statistics	
Multiple R	0.209674
R Square	0.043963
Adjusted R	0.024452

Square	
Standard Error	0.112204
Observation	101

ANOVA

	df	SS	MS	F	Significance F
Regression	2	0.056736	0.028368	2.253258	0.110472
Residual	98	1.233800	0.012589		
Total	100	1.290536			

	Coefficients	Standard Error	t Stat	p-value	Lower 95%	Upper 95%	Lower 90.000%	Upper 90.000%
Intercept	0.631109	0.030372	20.77905	1.09265E	0.570836	0.691382	0.580674	0.681544
AFQT	0.000842	0.000475	1.770944	0.079678	-0.000101	0.001785	5.24931E	0.001631
SELECT	0.046856	0.039192	1.195562	0.234752	-0.030918	0.124632	-0.018223	0.111937

SUMMARY

Regression Statistics	
Multiple R	0.278438
R Square	0.077527
Adjusted R Square	0.039091
Standard Error	0.111359
Observation	101

ANOVA

	df	SS	MS	F	Significance F
Regression	4	0.100052	0.025013	2.017047	0.098134
Residual	96	1.190483	0.012400		
Total	100	1.290536			

	Coefficients	Standard Error	t Stat	p-value	Lower 95%	Upper 95%	Lower 90.000%	Upper 90.000%
Intercept	0.633128	0.033045	19.15921	1.43449E	0.567533	0.698723	0.578243	0.688013
AFQT	0.000725	0.000480	1.511999	0.133818	-0.000227	0.001678	-7.14814E	0.001523
SCT	0.000796	0.000641	1.242406	0.217113	-0.000476	0.002069	-0.000268	0.001861
PMOS	-0.031393	0.022412	-1.400709	0.164525	-0.075882	0.013095	-0.068618	0.005831
SELECT	0.052482	0.039027	1.344753	0.181873	-0.024986	0.129950	-0.012337	0.117301

SUMMARY

Regression Statistics	
Multiple R	0.235712
R Square	0.055560
Adjusted R	0.026350

Square	
Standard Error	0.112095
Observation	101

ANOVA

	df	SS	MS	F	Significance F
Regression	3	0.071702	0.023900	1.902128	0.134327
Residual	97	1.218833	0.012565		
Total	100	1.290536			

	Coefficients	Standard Error	t Stat	p-value	Lower 95%	Upper 95%	Lower 90.000%	Upper 90.000%
Intercept	0.672016	0.020885	32.17627	1.47075E	0.630564	0.713468	0.637331	0.706701
SCT	0.000974	0.000634	1.536236	0.127735	-0.000284	0.002233	-7.89766E	0.002028
PMOS	-0.031750	0.022559	-1.407411	0.162501	-0.076524	0.013023	-0.069215	0.005714
SELECT	0.052045	0.039284	1.324841	0.188336	-0.025922	0.130013	-0.013194	0.117284

SUMMARY

Regression Statistics	
Multiple R	0.190457
R Square	0.036274
Adjusted R Square	0.016606
Standard Error	0.112654
Observation	101

ANOVA

	df	SS	MS	F	Significance F
Regression	2	0.046812	0.023406	1.844329	0.163577
Residual	98	1.243723	0.012691		
Total	100	1.290536			

	Coefficients	Standard Error	t Stat	p-value	Lower 95%	Upper 95%	Lower 90.000%	Upper 90.000%
Intercept	0.658592	0.018672	35.27008	1.70562E	0.621536	0.695647	0.627584	0.689599
SCT	0.000973	0.000637	1.526211	0.130178	-0.000292	0.002238	-8.56590E	0.002031
SELECT	0.047862	0.039367	1.215797	0.226982	-0.030260	0.125984	-0.017508	0.113233

SUMMARY

Regression Statistics	
Multiple R	0.253831
R Square	0.064430
Adjusted R Square	0.025448

Standard Error	0.112146
Observation	101

ANOVA

	df	SS	MS	F	Significance F
Regression	4	0.083149	0.020787	1.652823	0.167362
Residual	96	1.207386	0.012576		
Total	100	1.290536			

	Coefficients	Standard Error	t Stat	p-value	Lower 95%	Upper 95%	Lower 90.000%	Upper 90.000%
Intercept	0.603189	0.061711	9.774276	4.54271E	0.480692	0.725686	0.500693	0.705685
AFQT	0.000818	0.000505	1.619773	0.108561	-0.000184	0.001821	-2.07741E	0.001657
SCT	0.000742	0.000646	1.148312	0.253692	-0.000541	0.002026	-0.000331	0.001816
PMOS	-0.028762	0.022513	-1.277598	0.204472	-0.073450	0.015925	-0.066154	0.008628
MONTHS	0.000274	0.000413	0.662645	0.509145	-0.000547	0.001095	-0.000413	0.000961

SUMMARY

Regression Statistics	
Multiple R	0.227111
R Square	0.051579
Adjusted R Square	0.022247
Standard Error	0.112331
Observation	101

ANOVA

	df	SS	MS	F	Significance F
Regression	3	0.066565	0.022188	1.758447	0.160233
Residual	97	1.223970	0.012618		
Total	100	1.290536			

	Coefficients	Standard Error	t Stat	p-value	Lower 95%	Upper 95%	Lower 90.000%	Upper 90.000%
Intercept	0.610102	0.061518	9.917389	2.02164E	0.488005	0.732199	0.507937	0.712266
AFQT	0.000931	0.000496	1.875232	0.063768	-5.43666E	0.001916	0.000106	0.001755
PMOS	-0.028691	0.022549	-1.272349	0.206290	-0.073446	0.016063	-0.066140	0.008757
MONTHS	0.000304	0.000413	0.735481	0.463821	-0.000516	0.001125	-0.000382	0.000991

SUMMARY

Regression Statistics	
Multiple R	0.131137
R Square	0.017197
Adjusted R	-0.002860

Square	
Standard Error	0.113764
Observation	101

ANOVA

	df	SS	MS	F	Significance F
Regression	2	0.022193	0.011096	0.857404	0.427419
Residual	98	1.268342	0.012942		
Total	100	1.290536			

	Coefficients	Standard Error	t Stat	p-value	Lower 95%	Upper 95%	Lower 90.000%	Upper 90.000%
Intercept	0.688744	0.045582	15.10977	2.41777E	0.598287	0.779202	0.613052	0.764437
PMOS	-0.029423	0.022834	-1.288591	0.200574	-0.074737	0.015889	-0.067341	0.008493
MONTHS	8.35521E	0.000401	0.208016	0.835647	-0.000713	0.000880	-0.000583	0.000750

SUMMARY

Regression Statistics	
Multiple R	0.273082
R Square	0.074574
Adjusted R Square	0.036014
Standard Error	0.111537
Observation	101

ANOVA

	df	SS	MS	F	Significance F
Regression	4	0.096240	0.024060	1.934005	0.110974
Residual	96	1.194295	0.012440		
Total	100	1.290536			

	Coefficients	Standard Error	t Stat	p-value	Lower 95%	Upper 95%	Lower 90.000%	Upper 90.000%
Intercept	0.494590	0.084383	5.861200	6.46740E	0.327089	0.662090	0.354438	0.634741
AFQT	0.000927	0.000493	1.878837	0.063300	-5.23962E	0.001907	0.000107	0.001747
GDR	0.090519	0.057572	1.572257	0.119181	-0.023761	0.204799	-0.005102	0.186140
SELECT	0.045047	0.039086	1.152519	0.251969	-0.032537	0.122633	-0.019869	0.109965
MONTHS	0.000419	0.000414	1.013196	0.313513	-0.000402	0.001241	-0.000268	0.001107

SUMMARY

Regression Statistics	
Multiple R	0.201358
R Square	0.040545
Adjusted R	0.010871

Square	
Standard Error	0.112982
Observation	101

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	3	0.052324	0.017441	1.366357	0.257670
Residual	97	1.238211	0.012765		
Total	100	1.290536			

	Coefficients	*Standard Error*	*t Stat*	*p-value*	*Lower 95%*	*Upper 95%*	*Lower 90.000%*	*Upper 90.000%*
Intercept	0.568078	0.075740	7.500366	3.04470E	0.417755	0.718401	0.442295	0.693860
GDR	0.095074	0.058266	1.631698	0.105985	-0.020569	0.210717	-0.001690	0.191838
SELECT	0.042720	0.039572	1.079537	0.283025	-0.035820	0.121261	-0.022998	0.108439
MONTHS	0.000202	0.000402	0.503891	0.615479	-0.000596	0.001002	-0.000465	0.000871

SUMMARY

Regression Statistics	
Multiple R	0.195022
R Square	0.038033
Adjusted R Square	0.018401
Standard Error	0.112551
Observation	101

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	2	0.049083	0.024541	1.937332	0.149566
Residual	98	1.241452	0.012667		
Total	100	1.290536			

	Coefficients	*Standard Error*	*t Stat*	*p-value*	*Lower 95%*	*Upper 95%*	*Lower 90.000%*	*Upper 90.000%*
Intercept	0.593500	0.056275	10.54626	7.92172E	0.481822	0.705177	0.500051	0.686948
GDR	0.091213	0.057540	1.585193	0.116143	-0.022974	0.205400	-0.004336	0.186762
SELECT	0.041907	0.039389	1.063934	0.289973	-0.036258	0.120073	-0.023500	0.107314

SUMMARY

Regression Statistics	
Multiple R	0.234767
R Square	0.055115
Adjusted R Square	0.025892

Standard Error	0.112121
Observation	101

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	3	0.071128	0.023709	1.886027	0.137014
Residual	97	1.219407	0.012571		
Total	100	1.290536			

	Coefficients	*Standard Error*	*t Stat*	*p-value*	*Lower 95%*	*Upper 95%*	*Lower 90.000%*	*Upper 90.000%*
Intercept	0.569920	0.075135	7.585250	2.02003E	0.420797	0.719043	0.445142	0.694699
GDR	0.115730	0.058657	1.972982	0.051343	-0.000688	0.232150	0.018316	0.213144
PMOS	-0.037429	0.022867	-1.636817	0.104909	-0.082814	0.007955	-0.075405	0.000546
MONTHS	0.000188	0.000399	0.471347	0.638451	-0.000604	0.000980	-0.000475	0.000851

SUMMARY

Regression Statistics	
Multiple R	0.252686
R Square	0.063850
Adjusted R Square	0.034897
Standard Error	0.111601
Observation	101

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	3	0.082401	0.027467	2.205302	0.092354
Residual	97	1.208135	0.012455		
Total	100	1.290536			

	Coefficients	*Standard Error*	*t Stat*	*p-value*	*Lower 95%*	*Upper 95%*	*Lower 90.000%*	*Upper 90.000%*
Intercept	0.564782	0.058498	9.654706	7.46627E	0.448680	0.680885	0.467634	0.661931
SCT	0.001034	0.000632	1.635547	0.105175	-0.000220	0.002290	-1.59239E	0.002085
GDR	0.096604	0.057150	1.690366	0.094169	-0.016822	0.210032	0.001694	0.191515
SELECT	0.043788	0.039073	1.120660	0.265199	-0.033761	0.121338	-0.021101	0.108678

SUMMARY

Regression Statistics	
Multiple R	0.250390
R Square	0.062695
Adjusted R Square	0.033706

Standard Error	0.111670
Observation	101

ANOVA

	df	SS	MS	F	Significance F
Regression	3	0.080910	0.026970	2.162752	0.097356
Residual	97	1.209625	0.012470		
Total	100	1.290536			

	Coefficients	Standard Error	t Stat	p-value	Lower 95%	Upper 95%	Lower 90.000%	Upper 90.000%
Intercept	0.644764	0.031779	20.28887	1.12184E	0.581691	0.707837	0.591988	0.697540
AFQT	0.000835	0.000473	1.765338	0.080653	-0.000103	0.001774	4.95158E	0.001621
PMOS	-0.031292	0.022475	-1.392329	0.167006	-0.075900	0.013314	-0.068617	0.006032
SELECT	0.050968	0.039117	1.302958	0.195673	-0.026668	0.128605	-0.013994	0.115931

SUMMARY

Regression Statistics	
Multiple R	0.248534
R Square	0.061769
Adjusted R Square	0.032752
Standard Error	0.111725
Observation	101

ANOVA

	df	SS	MS	F	Significance F
Regression	3	0.079715	0.026571	2.128701	0.101550
Residual	97	1.210820	0.012482		
Total	100	1.290536			

	Coefficients	Standard Error	t Stat	p-value	Lower 95%	Upper 95%	Lower 90.000%	Upper 90.000%
Intercept	0.498357	0.084462	5.900308	5.31391E	0.330722	0.665993	0.358088	0.638626
AFQT	0.000909	0.000494	1.840127	0.068807	-7.14590E	0.001890	8.86659E	0.001730
GDR	0.094420	0.057570	1.640099	0.104223	-0.019839	0.208681	-0.001186	0.190028
MONTHS	0.000396	0.000414	0.957409	0.340740	-0.000425	0.001218	-0.000291	0.001084

SUMMARY

Regression Statistics	
Multiple R	0.245257
R Square	0.060151
Adjusted R Square	0.031083

Standard Error	0.111822
Observation	101

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	3	0.077627	0.025875	2.069363	0.109286
Residual	97	1.212909	0.012504		
Total	100	1.290536			

	Coefficients	*Standard Error*	*t Stat*	*p-value*	*Lower 95%*	*Upper 95%*	*Lower 90.000%*	*Upper 90.000%*
Intercept	0.637701	0.033006	19.32030	5.07995E	0.572192	0.703211	0.582886	0.692516
AFQT	0.000721	0.000482	1.495864	0.137935	-0.000235	0.001677	-7.94742E	0.001521
SCT	0.000769	0.000643	1.196037	0.234597	-0.000507	0.002047	-0.000299	0.001838
PMOS	-0.029116	0.022441	-1.297416	0.197564	-0.073656	0.015424	-0.066384	0.008152

APPENDIX K. CALCULATION FOR MALLOW'S COEFFICIENT

This appendix includes the information and calculations used to construct the C_p-versus-p plot. The F-statistics of the variables which were excluded from the model in question were obtained from Appendix J. A mnemonic for the variables in the model for which the C_p value is calculated are listed by the first letter of each variable included in the model. For example, a model with the "VARS IN" column listing ATGP has the variables **A**FQT, **T**est, **G**ender, and **P**MOS in the model equation. The other column headings are explained in Chapter III.

SUBSET#	IN VARS	P	M	F	P	CP
1	M	2	6	2.3493	2	8.7465
2	S	2	6	2.1702	2	7.8509
3	P	2	6	1.8956	2	6.4780
4	T	2	6	2.1518	2	7.7590
5	G	2	6	1.7143	2	5.5715
6	A	2	6	2.0085	2	7.0425
7	SM	3	6	2.5193	3	9.0771
8	PM	3	6	2.1483	3	7.5930
9	PS	3	6	2.0086	3	7.0344
10	TM	3	6	2.4263	3	8.7052
11	TS	3	6	2.2711	3	8.0844
12	TP	3	6	1.9340	3	6.7360
13	GM	3	6	2.0170	3	7.0680
14	GS	3	6	1.6528	3	5.6112
15	GP	3	6	1.6493	3	5.5972
16	TG	3	6	1.7800	3	6.1200
17	AM	3	6	2.4589	3	8.8356
18	AS	3	6	2.0997	3	7.3988
19	AP	3	6	1.7068	3	5.8272
20	AT	3	6	1.7814	3	6.1256
21	AG	3	6	1.4301	3	4.7204
22	PSM	4	6	2.4197	4	8.2591
23	TSM	4	6	2.7222	4	9.1666
24	TPM	4	6	2.2359	4	7.7077
25	TPS	4	6	2.1287	4	7.3861
26	GSM	4	6	2.0694	4	7.2082
27	GPM	4	6	2.0154	4	7.0462
28	GPS	4	6	1.6489	4	5.9467

29	TGM	4	6	2.1628	4	7.4883
30	TGS	4	6	1.7584	4	6.2752
31	TGP	4	6	1.7284	4	6.1852
32	ASM	4	6	2.7458	4	9.2374
33	APM	4	6	2.2053	4	7.6160
34	APS	4	6	1.8398	4	6.5194
35	ATM	4	6	2.3019	4	7.9057
36	ATS	4	6	1.8860	4	6.6580
37	ATP	4	6	1.3364	4	5.0092
38	AGM	4	6	1.9021	4	6.7063
39	AGP	4	6	1.2447	4	4.7341
40	AGS	4	6	1.3073	4	4.9219
41	ATG	4	6	1.1139	4	4.3417
42	TPSM	5	6	2.7371	5	8.4742
43	AGSM	5	6	1.9605	5	6.9210
44	**ATGP**	**5**	**6**	**0.7063**	**5**	**4.4126**
45	APSM	5	6	2.6730	5	8.3460
46	ATGM	5	6	1.6503	5	6.3006
47	GPSM	5	6	2.2467	5	7.4934
48	TGSM	5	6	2.3783	5	7.7566
49	TGPM	5	6	2.2533	5	7.5066
50	TGPS	5	6	1.8168	5	6.6336
51	ATSM	5	6	2.7397	5	8.4794
52	ATPM	5	6	1.9373	5	6.8746
53	ATPS	5	6	1.4644	5	5.9288
54	AGPM	5	6	1.8443	5	6.6886
55	AGPS	5	6	1.1143	5	5.2286
56	**ATGS**	**5**	**6**	**0.8574**	**5**	**4.7148**
57	AGPSM	6	6	2.1999	6	7.1999
58	AGTPM	6	6	1.3413	6	6.3413
59	ATPSM	6	6	2.7391	6	7.7391
60	**AGTPS**	**6**	**6**	**0.0540**	**6**	**5.0540**
61	AGTSM	6	6	1.6879	6	6.6879
62	GTPSM	6	6	3.0639	6	8.0639

APPENDIX L. VARIABLE INFLATION FACTORS

The variable inflation factors are listed in the table below. It can be seen that none of the VIF's approach the value five which is an indicatcr that multicollinearity is not a factor in the regression model.

VARIABLE	R2	VIF
AFQT	0.1191	1.1352
TEST	0.0431	1.0449
GENDER	0.0573	1.0608
PMOS	0.0368	1.0382
SELECTION	0.0115	1.0116
MONTH	0.0982	1.0888

The VIF is calculated using the coefficient of determination. These calculations are in the following pages.

w/o AFQT

Regression Statistics	
Multiple R	0.345058
R Square	0.119065
Adjusted R Square	0.072700
Standard Error	22.72502
Observation	101

VIF = 1.135158

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	5	6630.939	1326.187	2.568008	0.031764
Residual	95	49060.52	516.4265		
Total	100	55691.46			

	Coefficients	*Standard Error*	*t Stat*	*p-value*	*Lower 95%*	*Upper 95%*	*Lower 90.000%*	*Upper 90.000%*
Intercept	72.19629	15.65314	4.6:2255	1.24245E	41.12089	103.2716	46.19562	98.19696
GENDER	6.748161	11.92154	0.566047	0.572696	-16.91908	30.41540	-13.05413	26.55045
TEST	0.251706	0.128848	1.953502	0.053701	-0.004090	0.507503	0.037681	0.465730
PMOS	-1.165249	4.645174	-0.250851	0.802470	-10.38707	8.056580	-8.881120	6.550622
SELECT	-1.919172	7.980448	-0.240484	0.810473	-17.76235	13.92401	-15.17510	11.33676
MONTHS	-0.232983	0.081020	-2.875598	0.004977	-0.393829	-0.072136	-0.367562	-0.098403

w/o Months

Regression Statistics	
Multiple R	0.313353
R Square	0.098190
Adjusted R Square	0.050726
Standard Error	27.60087
Observation	101

VIF = 1.108881

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	5	7879.923	1575.984	2.068742	0.076056
Residual	95	72371.77	761.8082		
Total	100	80251.70			

	Coefficients	*Standard Error*	*t Stat*	*p-value*	*Lower 95%*	*Upper 95%*	*Lower 90.000%*	*Upper 90.000%*
Intercept	140.1180	15.35303	9.126409	1.21735E	109.6384	170.5976	114.6159	165.6202
AFQT	-0.343685	0.119518	-2.875598	0.004977	-0.580958	-0.106412	-0.542211	-0.145160
GENDER	-15.36743	14.41786	-1.065861	0.289187	-43.99049	13.25561	-39.31624	8.581367
TEST	0.083467	0.159376	0.523709	0.601699	-0.232935	0.399869	-0.181265	0.348199
PMOS	-0.013902	5.643704	-0.002463	0.998039	-11.21806	11.19026	-9.388384	9.360579
SELECT	-4.392914	9.685192	-0.453570	0.651172	-23.62043	14.83461	-20.48051	11.69468

w/o Gender

Regression Statistics	
Multiple R	0.239386
R Square	0.057305
Adjusted R Square	0.007690
Standard Error	0.195244
Observation	101

VIF = 1.060789

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	5	0.220145	0.044029	1.155000	0.337159
Residual	95	3.621438	0.038120		
Total	100	3.841584			

	Coefficients	*Standard Error*	*t Stat*	*p-value*	*Lower 95%*	*Upper 95%*	*Lower 90.000%*	*Upper 90.000%*
Intercept	0.997779	0.107966	9.241557	6.90833E	0.783438	1.212120	0.818441	1.177117
AFQT	0.000498	0.000879	0.566047	0.572696	-0.001248	0.002245	-0.000963	0.001959
TEST	-0.000757	0.001126	-0.672225	0.503071	-0.002993	0.001478	-0.002628	0.001113

PMOS	0.068269	0.039303	1.736994	0.085629	-0.009757	0.146297	0.002984	0.133555
SELECT	0.029809	0.068517	0.435060	0.664505	-0.106215	0.165833	-0.084001	0.143620
MONTHS	-0.000768	0.000721	-1.065861	0.289187	-0.002201	0.000663	-0.001967	0.000429

w/o Test

Regression Statistics	
Multiple R	0.207387
R Square	0.043009
Adjusted R Square	-0.007358
Standard Error	17.74231
Observation	101

VIF = 1.044942

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	5	1344.008	268.8017	0.853908	0.515169
Residual	95	29905.02	314.7897		
Total	100	31249.03			

	Coefficients	*Standard Error*	*t Stat*	*p-value*	*Lower 95%*	*Upper 95%*	*Lower 90.000%*	*Upper 90.000%*
Intercept	15.75873	13.42323	1.173988	0.243334	-10.88974	42.40721	-6.537953	38.05541
AFQT	0.153428	0.078540	1.953502	0.053701	-0.002493	0.309350	0.022969	0.283887
PMOS	0.594483	3.627357	0.163888	0.870166	-6.606726	7.795694	-5.430742	6.619710
SELECT	-1.504649	6.230633	-0.241492	0.809694	-13.87401	10.86471	-11.85404	8.844751
MONTHS	0.034489	0.065856	0.523709	0.601699	-0.096252	0.165231	-0.074901	0.143881
GENDER	-6.252506	9.301205	-0.672225	0.503071	-24.71772	12.21270	-21.70228	9.197271

w/o PMOS

Regression Statistics	
Multiple R	0.191865
R Square	0.036812
Adjusted R Square	-0.013881
Standard Error	0.501760
Observation	101

VIF = 1.038219

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	5	0.914115	0.182823	0.726168	0.605474
Residual	95	23.91756	0.251763		
Total	100	24 83168			

	Coefficients	*Standard Error*	*t Stat*	*p-value*	*Lower 95%*	*Upper 95%*	*Lower 90.000%*	*Upper 90.000%*
Intercept	0.015785	0.382355	0.041283	0.967156	-0.743286	0.774856	-0.619327	0.650897
AFQT	-0.000568	0.002264	-0.250851	0.802470	-0.005063	0.003927	-0.004329	0.003193
SELECT	0.112325	0.175882	0.638638	0.524594	-0.236844	0.461494	-0.179824	0.404474
MONTHS	-4.59451E	0.001865	-0.002463	0.998039	-0.003707	0.003698	-0.003102	0.003093
GENDER	0.450884	0.259577	1.736994	0.085629	-0.064441	0.966210	0.019713	0.882055
TEST	0.000475	0.002901	0.163888	0.870166	-0.005[illegible]	[illegible]06234	-0.004343	0.005294

w/o
Selection

Regression Statistics	
Multiple R	0.107195
R Square	0.011490
Adjusted R Square	-0.040535
Standard Error	0.292067
Observation	101

VIF = 1.011624

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	5	0.094202	0.018840	0.220864	0.952752
Residual	95	8.103817	0.085303		
Total	100	8.198019			

	Coefficients	*Standard Error*	*t Stat*	*p-value*	*Lower 95%*	*Upper 95%*	*Lower 90.000%*	*Upper 90.000%*
Intercept	0.088775	0.222379	0.399206	0.690637	-0.352702	0.530252	-0.280607	0.458157
AFQT	-0.000317	0.001318	-0.240484	0.810473	-0.002933	0.002299	-0.002506	0.001872
MONTHS	-0.000491	0.001084	-0.453570	0.651172	-0.002644	0.001661	-0.002293	0.001309
GENDER	0.066705	0.153324	0.435060	0.664505	-0.237681	0.371091	-0.187973	0.321384
TEST	-0.000407	0.001688	-0.241492	0.809694	-0.003759	0.002944	-0.003212	0.002396
PMOS	0.038058	0.059592	0.638638	0.524594	-0.080248	0.156365	-0.060928	0.137045

APPENDIX M: THE FORWARD REGRESSION MODEL

The forward regression model is developed by entering one variable at a time into an initial model. The initial model is a simple regression equation that includes the variable that has the largest correlation coefficient r. The variable AFQT (A) has the highest r. Therefore, the initial model is

$$MOE = \beta_0 + \beta_A X_A \text{ ; } r = 0.1733. \qquad \textbf{(Equation 27)}$$

<u>STEP 1</u>: To select a variable to enter the model, a variable is entered into the initial model, the MOE is regressed onto the variables, and the SSR is calculated. The variable would then be extracted and the next variable entered. The variable which produces the largest SSR is selected to enter the model. For this step the entering variable to join the variable **A** and the related SSR is listed.

VARIABLES	SSR
AM	0.04614
AP	0.05974
AS	0.05674
AG	***0.06827***
AT	0.05658

Entering the variable Gender (G) produced the largest SSR. It is selected as the next variable in the equation that now takes the form,

$$MOE = \beta_0 + \beta_A X_A + \beta_G X_G \text{ ; } r = 0.2300. \qquad \textbf{(Equation 28)}$$

<u>STEP 2</u>: The next step is to test the new model. The null hypothesis is that the coefficient of the entering variable is equal to zero. This hypothesis is

equivalent to the hypothesis that the reduced model is the true equation. The F-statistic is used for the test and is calculated with the equation

$$F = \frac{SSE_{reduced} - SSE_{full}}{MSE_{full}} = \frac{0.029533}{0.012472} = 2.3653 \qquad \textbf{(Equation 29)}$$

This value is compared to the critical F value at the 0.15 significance level (predetermined). The critical F value is $F_c = F_{1,98} = 2.0170$. The F-statistic is greater than F critical. Therefore, the null hypothesis is rejected and entrance of the variable **G** is permitted.

STEP 3: Steps 1 and 2 are iterated. The next variable selected to enter the equation was PMOS (P). The new equation is

$$MOE = \beta_0 + \beta_A X_A + \beta_G X_G + \beta_P X_P;\ r = 0.2787. \qquad \textbf{(Equation 30)}$$

The test of the null hypothesis produced a F-statistic with a value of 2.600 and a critical F value of 2.018. Again the null hypothesis is rejected and entrance of the variable **P** is allowed.

STEP 4: The next variable selected was Test (T). The F-statistic had a value of 1.8400 which was compared to the critical F value of 2.019. Since the F-statistic was too small to reject the null hypothesis, the iterations end. Variable **G** cannot enter the model, resulting in the final equation of

$$MOE = \beta_0 + \beta_A X_A + \beta_G X_G + \beta_P X_P, \qquad \textbf{(Equation 31)}$$

where,

$\beta_0 = 0.5555$

$\beta_A = 0.000759$

$\beta_G = 0.1047$

$\beta_P = -0.03646$

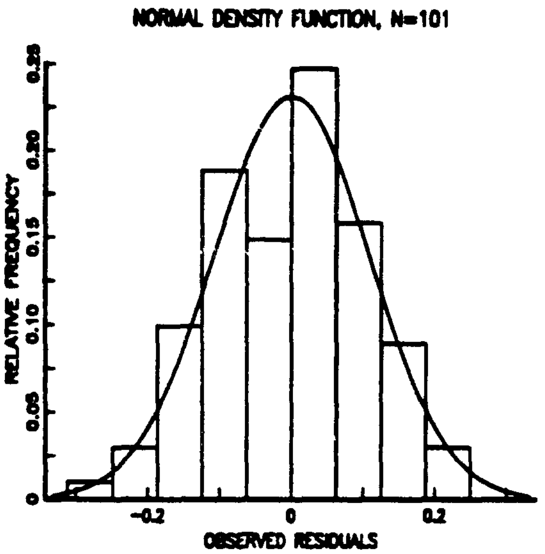
and **r = 0.2787**.

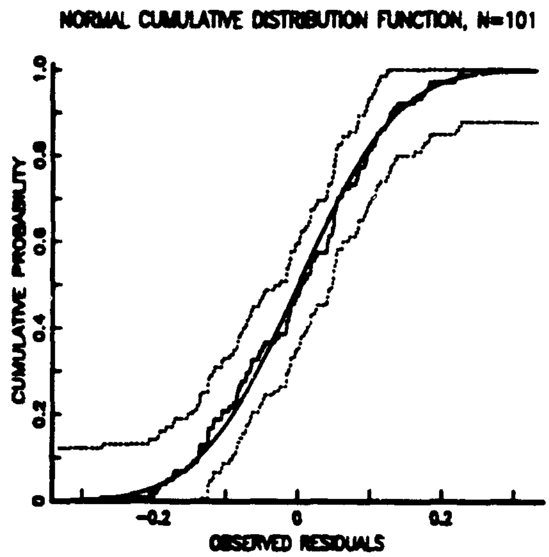
A forward regression model was calculated for a significance level of 0.10, but only the variable **A** was able to enter the model. The F-statistic for the next entering variable was 2.3653 which was compared to the critical F value of 2.7650. Since the F-statistic was not large enough to reject the null hypothesis, the iterations ended and the final forward regression model was

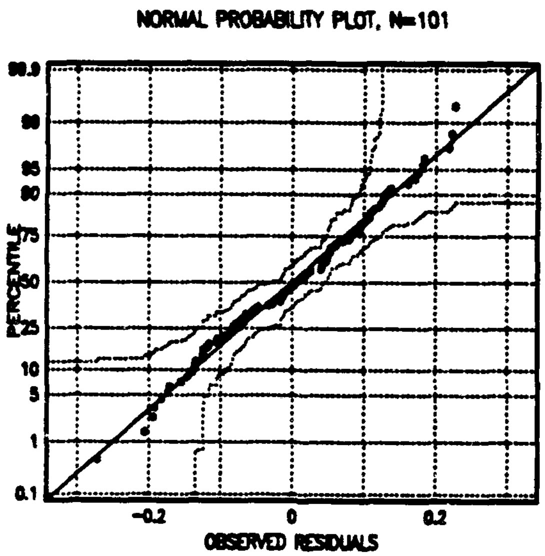
$$MOE = \beta_0 + \beta_A X_A \qquad \textbf{(Equation 32)}$$

with an **r = 0.1733**.

APPENDIX N. GRAPHICAL ANALYSIS OF RESIDUALS FOR NORMALITY







Three graphical methods are presented in the above figures to help determine the distribution of the regression model's residuals. The upper left is a frequency histogram, the upper right is a normal cumulative distribution, and the lower graph is a quantile-quantile plot of a theoretical and empirical normal distribution. The Kolmogorov-Smirnov bounds in both the normal cumulative distribution and quantile-quantile plots have a 90% confidence interval.

APPENDIX O. STANDARDIZED RESIDUAL CALCULATIONS

This appendix includes the spread sheet used to calculate the standard residuals from the observed residuals of the model. Each standardized residual was checked to determine if it was within ± 2. A binary was used to indicate if the value met this criterion, or not. The percentage of the values that met the criterion was calculated and compared with the target value of 95%, described in Chapter IV.

Observation	Predicted SUCC	Residuals	Standard Residuals	W/I 2
1	0.653703619	0.12407438	1.124884681	1
2	0.738797923	0.00231344	0.020974204	1
3	0.745070456	0.10715191	0.971462002	1
4	0.752796501	0.13720360	1.243917048	1
5	0.725801551	0.05308755	0.481303079	1
6	0.693539866	-0.06465060	-0.586136124	1
7	0.690220564	0.01311306	0.118885843	1
8	0.704986719	-0.01081982	-0.098094867	1
9	0.703490591	0.00095414	0.008650518	1
10	0.717400345	0.02259991	0.204895625	1
11	0.658071337	0.12081777	1.095359585	1
12	0.711078013	-0.08107764	-0.735067123	1
13	0.653448946	0.05099579	0.462338696	1
14	0.642164351	-0.08549757	-0.775139132	1
15	0.719255362	0.04130026	0.374436999	1
16	0.587400801	0.04148845	0.37614318	1
17	0.646343646	-0.01634327	-0.148171607	1
18	0.649392884	-0.20605966	-1.868180666	1
19	0.726167623	0.05272148	0.477984198	1
20	0.666861188	-0.05519429	-0.500403228	1
21	0.707893742	0.07099536	0.64365908	1
22	0.717894266	0.17099473	1.550274566	1
23	0.732301533	-0.12436463	-1.127516168	1
24	0.685039062	-0.08948343	-0.811275793	1
25	0.662208015	0.04001450	0.362779997	1
26	0.727390095	-0.13516813	-1.225463014	1
27	0.69103794	0.16118443	1.461332272	1
28	0.641702272	-0.01281301	-0.11616549	1
29	0.691578068	-0.13602206	-1.233204953	1
30	0.714513365	0.04604226	0.417428954	1
31	0.734121116	-0.06745411	-0.611553337	1
32	0.720218544	0.09422608	0.854273701	1
33	0.671314139	0.18090823	1.640152438	1
34	0.618987325	0.09167956	0.831186402	1
35	0.561130811	-0.11779759	-1.067977971	1

36	0.657477818	-0.13970022	-1.266551927	1
37	0.655071898	0.04937284	0.447624669	1
38	0.657820785	0.22648484	2.053359695	0
39	0.707897334	-0.07789696	-0.706230412	1
40	0.69940012	0.06115551	0.554448835	1
41	0.672960195	-0.00518241	-0.046984872	1
42	0.68340431	0.13215143	1.198112924	1
43	0.676021763	0.17508934	1.587397213	1
44	0.661795736	-0.11651825	-1.056379249	1
45	0.638478977	-0.04736812	-0.429449496	1
46	0.716988066	-0.12476543	-1.131149936	1
47	0.669386749	0.18283562	1.657626564	1
48	0.677660106	0.04970115	0.45060122	1
49	0.644524063	-0.05341321	-0.484255527	1
50	0.726622519	-0.14598784	-1.323556834	1
51	0.683358102	-0.01558032	-0.14125451	1
52	0.662613112	0.07960903	0.721752445	1
53	0.649893987	0.12788401	1.159423613	1
54	0.703603049	0.02592103	0.235005569	1
55	0.70307472	0.22248091	2.017059169	0
56	0.658317767	0.21793223	1.975819897	1
57	0.70696856	-0.00363493	-0.032955043	1
58	0.635756785	-0.15464850	-1.402076185	1
59	0.720939913	0.04767134	0.432198555	1
60	0.665612551	-0.07227914	-0.655298037	1
61	0.648860963	0.05447266	0.493860762	1
62	0.678562714	0.06143754	0.557005835	1
63	0.64462007	-0.03295318	-0.298760524	1
64	0.725755343	0.10924465	0.990435253	1
65	0.658071337	0.08304003	0.752858572	1
66	0.700831058	0.11361357	1.030044767	1
67	0.717123098	-0.10545620	-0.956088376	1
68	0.631570308	-0.01184778	-0.107414557	1
69	0.697966651	-0.10574402	-0.958697752	1
70	0.737301795	0.00380957	0.034538418	1
71	0.742980809	0.01757482	0.159337059	1
72	0.717666818	-0.04988903	-0.452304614	1
73	0.675698307	-0.08458745	-0.766887851	1
74	0.606481283	-0.01537043	-0.13935161	1
75	0.658472842	0.00819415	0.074289977	1
76	0.686087477	0.01724615	0.156357287	1
77	0.707620086	-0.19396913	-1.758565378	1
78	0.689334406	0.12622133	1.144349414	1
79	0.695984811	0.08290429	0.75162798	1
80	0.693539866	-0.02687286	-0.243635112	1
81	0.690352004	-0.17146296	-1.554519644	1
82	0.694950761	-0.12453387	-1.129050521	1
83	0.681789601	0.09487728	0.860177648	1
84	0.713483932	0.10207180	0.925404684	1
85	0.661387048	-0.08291452	-0.751720717	1
86	0.694943579	-0.06494320	-0.588788919	1
87	0.716329356	-0.12410672	-1.125177934	1
88	0.69037923	-0.17371275	-1.574916676	1

89	0.679881194	-0.27210382	-2.466951043	0
90	0.722713288	0.08125516	0.736676551	1
91	0.653907433	0.04942619	0.448108395	1
92	0.727709959	-0.13659910	-1.238436524	1
93	0.68106105	0.09671695	0.87685641	1
94	0.701513402	0.03959796	0.359003588	1
95	0.674657075	-0.09618455	-0.872029601	1
96	0.66879323	0.10787366	0.978005511	1
97	0.708352229	-0.19168574	-1.737863706	1
98	0.685995061	0.01733856	0.157195147	1
99	0.701496951	-0.18260791	-1.655562093	1
100	0.626804676	-0.07124867	-0.645955628	1
101	0.678335266	-0.01055748	-0.095716413	1
				0.970297
			95% <	**97%**

APPENDIX P: CHI-SQUARE GOODNESS-OF-FIT CALCULATIONS

```
                    ANALYSIS OF NORMAL DISTRIBUTION FIT

DATA         :  ACTRES
SELECTION    :  ALL
X AXIS LABEL:  OBSERVED RESIDUALS
SAMPLE SIZE :  101
CENSORING    :  NONE
FREQUENCIES :  1
EST. METHOD :  MAXIMUM LIKELIHOOD
CONF METHOD :  EXACT

                        CONF. INTERVALS        COVARIANCE MATRIX OF
                         (90 PERCENT)          PARAMETFR ESTIMATES
PARAMETER  ESTIMATE    LOWER      UPPER        MU          SIGMA
MU         2.2772E-7 -0.017697 0.017697        0.00011449 0
SIGMA      1.0754E-1  0.096918 0.12242         0           0.000057246
LOG LIKELIHOOD FUNCTION AT MLE = 81.911

            SAMPLE       FITTED                     GOODNESS OF FIT TESTS
MEAN     :  2.2772E-7    2.2772E-7              CHI-SQUARE :  3.9781
STD DEV :   1.0807E-1    1.0754E-1                DEG FREED:  5
SKEWNESS:  -8.1138E-2    0.0000E0                 SIGNIF   :  0.55257
KURTOSIS:   2.3985E0     3.0000E0               KOLM-SMIRN :  0.069393
• BASED ON MIDPOINTS OF FINITE INTERVALS          SIGNIF   :  0.71558
                                                CRAMER-V M :  0.066142
PERCENTILES   SAMPLE       FITTED                 SIGNIF   :   > .15
     5:      -0.17371    -1.7092E-1             ANDER-DARL :  0.37935
    10:      -0.1366     -1.3783E-1               SIGNIF   :   > .15
    25:      -0.08291    -7.2499E-2
    50:       0.00381     2.3858E-7             KS, AD, AND CV SIGNIF. LEVELS NOT
    75:       0.081255    7.2499E-2             EXACT WITH ESTIMATED PARAMETERS.
    90:       0.12788     1.3783E-1
    95:       0.17509     1.7092E-1             NOTE:  A SMALL SIGNIFICANCE LEVEL
                                                (EG. P<.01) INDICATES LACK OF FIT

                  CHI-SQUARE GOODNESS OF FIT TABLE

  LOWER      UPPER      OBS    EXP       O-E       ((O-E)•2)÷E
  -INF.     -0.18697     4     4.1455  -0.14551    0.0051073
-0.18697    -0.12465    10     8.2979   1.7021     0.34814
-0.12465     0.062323   19    15.948    3.0519     0.58403
-0.062323    0          15    22.108   -7.1084     2.2855
 0           0.062323   25    22.108    2.8916     0.37819
 0.062323    0.12465    16    15.948    0.051845   0.00016854
 0.12465     0.18697     9     8.298    0.70205    0.059397
 0.18697    +INF.        3     4.1455  -1.1455     0.31655
TOTAL                  101   101                   3.9781
```

The Chi-Square Goodness-of-Fit calculations are shown in the above table. The expected and observed frequency figures are displayed and the ensuing Chi-Square conversion for each interval is shown to produce the resulting Chi-Square statistic.

APPENDIX Q. MEAN SQUARE PREDICTED RESIDUAL CALCULATIONS

This appendix contains the calculations of the MSPR from the validation data set when the ATGP model from the model-building set is applied. The individual vectors of variable measurements for each recruiter is presented in the order that the ATGP model can use it. The column headings are self-explanatory. A comparison is presented at the end of the calculations.

AFQT	TEST	GENDER	PMOS	SELECT	MONTHS	SUCC RATE	PRED SUCC RATE	SQR ERR
52	20	1	1	0	132	0.54527748	0.685492	0.0196601
52	-7	1	1	1	96	0.59111085	0.67021	0.0062566
66	42	1	0	0	147	0.58063467	0.739198	0.0251423
36	14	1	0	0	141	0.66777778	0.70202	0.0011725
99	48	1	0	0	122	0.89000011	0.766057	0.0153618
33	28	1	0	1	109	0.62888926	0.707811	0.0062286
42	49	1	0	0	28	0.74000026	0.726096	0.0001933
38	26	1	1	0	125	0.77888911	0.678934	0.0099910
85	48	1	1	0	72	0.70333363	0.724803	0.0004609
58	20	1	1	0	115	0.59333341	0.689758	0.0092977
77	21	1	1	0	123	0.74000026	0.703833	0.0013080
80	10	1	0	1	74	0.77888911	0.73104	0.0022895
89	44	1	0	0	136	0.76055563	0.756683	1.49973E-
75	25	1	0	0	130	0.66777778	0.735975	0.0046508
82	14	1	1	1	70	0.59111085	0.703426	0.0126146
79	32	1	1	0	108	0.81555574	0.711481	0.0108315
66	0	1	0	0	100	0.51888904	0.715426	0.0386267
23	37	1	0	0	138	0.57041689	0.705795	0.0183272
50	21	1	1	0	88	0.57847252	0.684636	0.0112706
61	9	1	0	0	94	0.63000037	0.716[illegible]	0.0075628
97	50	1	1	0	69	0.59222263	0.734467	0.0202334
75	24	1	1	0	53	0.40777737	0.704109	0.0878124
72	-3	1	0	0	80	0.555556	0.717994	0.0263861
25	-12	1	0	0	116	0.70333363	0.679483	0.0005688
84	30	1	0	0	149	0.59111085	0.745204	0.0237446
90	38	1	1	0	49	0.74111137	0.722698	0.0003390
52	15	1	1	0	94	0.51777759	0.682662	0.0271868
61	24	1	0	0	43	0.63000037	0.725455	0.0091115
16	6	1	1	0	145	0.555556	0.651972	0.0092960
78	20	1	1	0	126	0.66777778	0.703978	0.0013104
								0.4072519
							MSPR =	0.0135750
							MSE =	0.01219
							% DIFFERENCE	11.362313

APPENDIX R. MATRIX CALCULATIONS

The matrix calculations presented in this appendix are those operations that resulted in the inverted matrix used in the predictive interval in Chapter IV. The calculations started with the full database of 101 recruiters. Those variables which were included in the final model comprised the ***X*** matrix. The natural layout of the spreadsheet eased the manipulation, but the matrix functions embedded in the computer program made calculations even easier. The steps to calculating inverse of the $\boldsymbol{X^T X}$ matrix are shown in sequence.

X

CONST	AFQT	TEST	GENDER	PMOS
1	61	4	1	1
1	54	35	1	1
1	96	34	1	0
1	52	20	1	1
1	95	42	1	0
1	52	-7	1	1
1	59	36	1	0
1	49	31	1	1
1	43	45	1	1
1	52	0	1	1
1	66	42	1	0
1	36	14	1	0
1	56	18	1	1
1	17	32	1	1
1	99	48	1	0
1	81	30	1	0
1	33	28	1	0
1	40	19	1	0
1	70	14	1	0
1	96	36	1	1
1	50	27	1	0
1	86	0	1	0
1	42	49	1	0
1	17	-1	1	0
1	38	26	1	1
1	85	48	1	1
1	26	9	1	1
1	72	31	1	0
1	58	20	1	1
1	66	24	1	0
1	32	-23	1	0
1	77	21	1	1
1	46	-28	1	0
1	32	12	1	1
1	25	20	1	1
1	85	27	1	0
1	87	18	1	0
1	36	31	0	0

X

CONST	AFQT	TEST	GENDER	PMOS
1	74	26	1	0
1	89	44	1	0
1	75	25	1	0
1	79	39	1	0
1	40	13	1	0
1	82	14	1	1
1	85	17	0	0
1	35	33	1	1
1	78	-3	1	1
1	93	23	1	0
1	24	26	1	0
1	85	6	1	0
1	79	32	1	1
1	83	-6	1	0
1	33	28	1	0
1	66	0	1	0
1	23	37	1	0
1	78	24	1	1
1	82	15	1	0
1	50	21	1	1
1	61	9	1	0
1	44	17	1	0
1	97	50	1	1
1	82	31	1	1
1	75	24	1	1
1	72	-18	1	1
1	68	36	1	0
1	72	-3	1	0
1	86	56	1	1
1	25	-12	1	0
1	71	47	1	0
1	41	53	1	0
1	84	30	1	0
1	32	3	1	0
1	49	58	0	0
1	85	18	1	1
1	90	38	1	1
1	30	5	0	0

1	38	26	1	1	1	60	29	1	1
1	44	8	1	1	1	52	15	1	1
1	93	35	1	1	1	36	24	1	1
1	23	27	1	1	1	63	20	1	1
1	66	31	1	0	1	41	24	1	1
1	68	40	1	0	1	59	26	1	0
1	25	3	1	0	1	32	20	1	0
1	52	-15	1	1	1	17	49	1	0
1	80	10	1	0	1	16	6	1	1
1	98	28	1	1	1	49	23	1	0
1	76	47	1	0	1	61	24	1	0

X

CONST	AFQT	TEST	GENDER	PMOS
1	78	20	1	1
1	20	57	1	1
1	32	17	1	0

XTX

101	5943	2281	97	44
5943	405387	141892.5	5743	2576
2281	141892.5	82763.5	2170	993
97	5743	2170	97	44
44	2576	993	44	44

(XTX)-1

0.309367	-0.000805	-0.000687	-0.245949	-0.000726
-0.000805	1.87482E	-4.74236E	-0.000209	2.43817E
-0.000687	-4.74236E	3.33180E	0.000230	-1.77504E
-0.245949	-0.000209	0.000230	0.272188	-0.019204
-0.000726	2.43817E	-1.77504E	-0.019204	0.041631

APPENDIX S. PREDICTED SUCCESS RATE CALCULATIONS FOR BNCOC GROUP

This appendix includes the database of 46 BNCOC students discussed in Chapters I and IV. Using the final ATGP model, apredicted success rate was calculated for each student. From the PI described in Chapter IV a PI was computed for each student. The lower bound of the PI is recorded in the spreadsheet shown in this appendix. This lower bound value is the predicted success rate considered when determining the student's entry into recruiting.

CNST	AFQT	TEST	GENDER	PMOS	PRED RATE
1	89	34	1	0	0.7342753
1	96	47	1	1	0.7130348
1	96	13	1	0	0.7205924
1	87	33	1	1	0.6952195
1	87	23	1	1	0.6865835
1	64	17	1	0	0.7036916
1	71	10	1	0	0.7020991
1	65	14	1	0	0.7017369
1	95	13	1	1	0.6830363
1	89	18	1	1	0.6835377
1	48	27	1	0	0.70215
1	61	11	1	0	0.6966017
1	80	18	1	1	0.6778128
1	66	1	1	0	0.6911462
1	87	8	1	1	0.6736295
1	69	22	1	1	0.6742701
1	79	12	1	1	0.6719951
1	47	5	1	0	0.6825147
1	96	-14	1	1	0.6603552
1	95	-14	1	1	0.6597191
1	73	4	1	1	0.6612697
1	99	-24	1	1	0.6536275
1	76	-6	1	1	0.654542
1	91	-22	1	1	0.6502659
1	44	19	1	1	0.6557768
1	37	-3	1	0	0.6692449
1	72	-9	1	1	0.6494068
1	59	0	1	1	0.6489099
1	89	-28	1	1	0.6438121
1	33	20	1	1	0.6496433
1	50	-2	1	1	0.6414578
1	28	16	1	1	0.6430084
1	52	-7	1	1	0.638412
1	35	-4	1	1	0.6301891
1	41	-10	1	1	0.6288241
1	16	-1	1	1	0.620694
1	44	-28	1	1	0.6151876

1	43	-29	1	1	0.6136879
1	44	-37	1	1	0.6074152
1	26	-22	1	1	0.6089194
1	14	-14	1	1	0.608195
1	35	-35	1	1	0.6034175
1	34	22	0	0	0.5783266
1	56	-78	1	1	0.5796408
1	62	-8	0	0	0.5702294
1	53	-90	1	1	0.5673693

(XTX)-1

0.309367	-0.000806	-0.00069	-0.245950	-0.000727
-0.000806	0.000019	-0.00000	-0.000209	0.000024
-0.000687	-0.000005	0.000033	0.000231	-0.000018
-0.245950	-0.000209	0.000231	0.272188	-0.019204
-0.000727	0.000024	-0.00002	-0.019204	0.041631

aXXa	LOWPI
0.036265	0.589214
0.060236	0.566305
0.049487	0.574608
0.038717	0.549987
0.037796	0.541415
0.020319	0.559751
0.027505	0.557652
0.022036	0.557675
0.054000	0.536739
0.042130	0.538066
0.022757	0.558037
0.023214	0.552457
0.032981	0.532981
0.035840	0.546115
0.048909	0.527686
0.024843	0.530010
0.036342	0.526928
0.029701	0.537913
0.106572	0.510454
0.104840	0.509935
0.040677	0.515900
0.143531	0.501243
0.060360	0.507804
0.122374	0.499298
0.026625	0.511392
0.044305	0.523622
0.063352	0.502462
0.039798	0.503601
0.139921	0.491668
0.034559	0.504701
0.042215	0.495981
0.039754	0.497703
0.050824	0.492335
0.050705	0.484120
0.058418	0.482220
0.065659	0.473590
0.104916	0.465398
0.108256	0.463672
0.136700	0.455487
0.095008	0.459803
0.089033	0.459486
0.130683	0.451892

aXX

-0.031672	0.000254	0.015473	-0.018364
-0.046976	0.000636	-0.002192	0.023206
-0.022876	-0.000478	0.009160	-0.017820
-0.030100	0.000212	-0.003542	0.023235
-0.023225	-0.000120	-0.005852	0.023413
0.000159	-0.000193	0.016777	-0.018671
-0.000669	-0.000460	0.013696	-0.018377
0.001415	-0.000298	0.015875	-0.018594
-0.022797	-0.000491	-0.009834	0.023785
-0.021400	-0.000296	-0.007424	0.023550
0.006177	0.000215	0.022433	-0.019239
0.006701	-0.000379	0.016019	-0.018638
-0.014147	-0.000253	-0.005542	0.023331
0.009546	-0.000736	0.012664	-0.018339
-0.012913	-0.000620	-0.009315	0.023679
-0.008033	-6.85099E	-0.002317	0.022992
-0.009217	-0.000449	-0.006718	0.023413
0.022107	-0.000512	0.017562	-0.018873
-0.005042	-0.001396	-0.016278	0.024289
-0.004236	-0.001391	-0.016069	0.024265
0.001117	-0.000687	-0.007311	0.023409
-0.000584	-0.001743	-0.019215	0.024540
0.005574	-0.001034	-0.010247	0.023659
0.004486	-0.001638	-0.017080	0.024309
0.014173	-4.99050E	0.002218	0.022435
0.035665	-0.000731	0.017806	-0.018975
0.010860	-0.001115	-0.010104	0.023615
0.015148	-0.000754	-0.005306	0.023138
0.010222	-0.001829	-0.018047	0.024367
0.022350	3.55790E	0.004750	0.022150
0.023775	-0.000778	-0.003885	0.022955
0.029129	-7.39813E	0.004872	0.022099
0.025601	-0.000954	-0.005458	0.023092
0.037237	-0.000773	-0.001210	0.022624
0.036527	-0.001001	-0.003850	0.022877
0.050485	-0.000583	0.003456	0.022108
0.046484	-0.001615	-0.008634	0.023270
0.047977	-0.001644	-0.008656	0.023263
0.052671	-0.001915	-0.010713	0.023429
0.056864	-0.001330	-0.003484	0.022724
0.061034	-0.001007	0.000872	0.022290
0.058548	-0.001806	-0.008368	0.023175
0.266845	-0.000115	-0.247981	-0.000288
0.071187	-0.003338	-0.022691	0.024450
0.264906	-0.001248	-0.260765	0.000927
0.081854	-0.003724	-0.024834	0.024590

0.255028	0.418686
0.357397	0.413617
0.299351	0.407794
0.439575	0.396394

APPENDIX T. THE TAYLOR AND RUSSELL TABLES

This appendix provides a copy of the completeTaylor and Russell tables. These tables are useful for determining the number of candidates to choose from a group of applicants.

H.C.TAYLOR and J.T. RUSSELL

Proportion of Employees Considered Satisfactory =.20

Selection Ratio

r	*.05*	*.10*	*.20*	*.30*	*.40*	*.50*	*.60*	*.70*	*.80*	*.90*	*.95*
.00	.20	.20	.20	.20	.20	.20	.20	.20	.20	.20	.20
.05	23	.23	.22	.22	.21	.21	21	.21	.20	.20	.20
.10	.26	.25	.24	.23	.23	.22	.22	.21	.21	.21	.20
.15	.30	.28	.26	.25	.24	.23	.23	.22	.21	.21	.20
.20	.33	.31	.28	.27	.26	.25	.24	.23	.22	.21	.21
.25	.37	.34	.31	.29	.27	.26	.24	.23	.22	.21	.21
.30	.41	.37	.33	.30	.28	.27	.25	.24	.23	.21	.21
.35	.35	.45	.41	.36	.32	.30	.28	.26	.24	.22	21
.40	.49	.44	.38	.34	.31	.29	.27	.25	.23	.22	.21
.45	.54	.48	.41	.36	.33	.30	.28	.26	.24	.22	.21
.50	.59	.52	.44	.38	.35	.31	.29	.26	.24	.22	.21
.55	.63	.56	.47	.41	.36	.32	.29	.27	.24	.22	.21
.60	.68	.60	.50	.43	.38	.34	.30	.27	.24	.22	.21
.65	.73	.64	.53	.45	.39	.35	.31	.27	.25	.22	.21
.70	.79	.69	.56	.48	.41	.36	.31	.28	.25	.22	.21
.75	.84	.74	.60	.50	.43	.37	.32	.28	.25	.22	.21
.80	.89	.79	.64	.53	.45	.38	.33	.28	.25	.22	.21
.85	.94	.85	.69	.56	.47	.39	.33	.28	.25	.22	.21
.90	.98	.91	.75	.60	.48	.40	.33	.29	.25	.22	.21
.95	1.00	.97	.82	.64	.50	.40	.33	.29	.25	.22	.21
1.00	1.00	1.00	1.00	.67	.50	.40	.33	.29	.25	.22	.21

Proportion of Employees Considered Satisfactory =.30

Selection Ratio

r	*.05*	*.10*	*.20*	*.30*	*.40*	*.50*	*.60*	*.70*	*.80*	*.90*	*.95*
.00	.30	.30	.30	.30	.30	.30	.30	.30	.30	.30	.30
.05	.34	.33	.33	.32	.32	.31	.31	.31	.31	30	.30
.10	.38	.36	.35	.34	.33	.33	.32	.32	.31	.31	.30
.15	.42	.40	.38	.36	.35	.34	.33	.33	.32	.31	.31
.20	.46	.43	.40	.38	.37	.36	.34	.33	.32	.31	.31
.25	.50	.47	.43	.41	.39	.37	.36	.34	.33	.32	.31
.30	.54	.50	.46	.43	.40	.38	.37	.35	.33	.32	.31
.35	.58	.54	.49	.45	.42	.40	.38	.36	.34	.32	.31
.40	.63	.58	.51	.47	.44	.41	.39	.37	.34	.32	.31
.45	.67	.61	.55	.50	.46	.43	.40	.37	.35	.32	.31
.50	.72	.65	.58	.52	.48	.44	.41	.38	.35	.33	.31
.55	.76	.69	.61	.55	.50	.46	.42	.39	.36	.33	.31
.60	.81	74	.64	.58	.52	47	.43	.40	.36	.33	.31
.65	.85	.78	.68	.60	.54	.49	.44	.40	.37	.33	.32
.70	.89	.82	.72	.63	.57	.51	.46	.41	.37	.33	.32
.75	.93	.86	.76	.67	.59	.52	.47	.42	.37	.33	.32
.80	96	90	.80	.70	.62	.54	.48	.42	.37	.33	.32
.85	.99	.94	.85	.74	.65	.56	.49	.43	.37	.33	.32
.90	1.00	.98	.90	.79	.68	.58	.49	.43	.37	.33	.32
.95	1.00	1.00	96	.85	.72	.60	.50	.43	.37	.33	.32
1.00	1.00	1.00	1.00	1.00	.75	.60	.50	.43	.38	.33	.32

Proportion of Employees Considered Satisfactory =.40

Selection Ratio

r	*.05*	*.10*	*.20*	*.30*	*.40*	*.50*	*.60*	*.70*	*.80*	*.90*	*.95*
.00	.40	.40	.40	.40	.40	.40	.40	.40	.40	.40	.40
.05	.44	.43	.43	.42	.42	.42	.41	.41	.41	.40	.40
.10	.48	.47	.46	.45	.44	.43	.42	.42	.41	.41	.40
.15	.52	.50	.48	.47	.46	.45	.44	.43	.42	.41	.41
.20	.57	.54	.51	.49	.48	.46	.45	.44	.43	.41	.41
.25	.61	.58	.54	.51	.49	.48	.46	.45	.43	42	.41
.30	.65	.61	.57	.54	.51	.49	.47	.46	.44	.42	.41
.35	.69	.65	60	.56	.53	.51	.49	.47	.45	42	.41
.40	.73	.69	.63	.59	.56	.53	.50	.48	.45	.43	.41
.45	.77	.72	.66	.61	.58	.54	.51	.49	.46	.43	.42
.50	.81	.76	.69	.64	.60	.56	.53	.49	.46	.43	42
.55	.85	.79	.72	.67	.62	.58	.54	.50	.47	.44	.42
.60	.89	.83	.75	.69	.64	.60	.55	.51	.48	.44	.42
.65	.92	.87	.79	.72	.67	.62	.57	.52	.48	.44	.42
.70	.95	.90	.82	.76	.69	.64	.58	.53	.49	.44	.42
.75	.97	.93	.86	.79	.72	.66	.60	.54	.49	.44	.42
.80	.99	.96	.89	.82	.75	.68	.61	.55	.49	.44	.42
.85	1.00	.98	.93	.86	.79	.71	.63	.56	.50	.44	.42
.90	1.00	1.00	.97	.91	.82	.74	.65	.57	.50	.44	.42
.95	1.00	1.00	.99	.96	.87	.77	.66	.57	.50	.44	.42
1.00	1.00	1.00	1.00	1.00	1.00	.80	.67	.57	.50	.44	.42

Proportion of Employees Considered Satisfactory =.50

Selection Ratio

r	*.05*	*.10*	*.20*	*.30*	*.40*	*.50*	*.60*	*.70*	*.80*	*.90*	*.95*
.00	.50	50	.50	.50	.50	50	.50	50	.50	.50	50
.05	.54	.54	.53	.52	.52	.52	51	.51	.51	.50	.50
.10	.58	.57	.56	.55	.54	.53	.53	.52	.51	.51	.50
.15	.63	.61	.58	.57	.56	.55	.54	.53	.52	.51	.51
.20	.67	.64	.61	.59	.58	.56	.55	.54	.53	.52	.51
.25	.70	.67	.64	.62	.60	.58	.56	.55	.54	.52	.51
.30	.74	.71	.67	.64	.62	60	.58	.56	.54	.52	.51
.35	.78	.74	.70	.66	.64	.61	.59	.57	.55	.53	.51
.40	.82	.78	.73	.69	.66	.63	.61	.58	.56	.53	.52
.45	.85	.81	.75	.71	.68	.65	.62	.59	.56	.53	.52
.50	.88	.84	.78	.74	.70	.67	.63	.60	.57	.54	.52
.55	.91	.87	.81	.76	.72	.69	.65	.61	.58	.54	.52
.60	.94	.90	.84	.79	.75	.70	.66	.62	.59	.54	52
.65	.96	.92	.87	.82	.77	.73	.68	.64	.59	.55	.52
.70	.98	.95	.90	.85	.80	.75	.70	.65	.60	.55	.53
.75	.99	.97	.92	.87	.82	.77	.72	.66	.61	.55	.53
.80	1.00	.99	.95	.90	.85	.80	.73	.67	.61	.55	.53
.85	1.00	.99	.97	.94	.88	.82	.76	.69	.62	.55	.53
.90	1.00	1.00	.99	.97	.92	.86	.78	7.0	.62	.56	.53
.95	1.00	1.00	1.00	.99	.96	.90	.81	.71	.63	.56	.53
1.00	1.00	1.00	1.00	1.00	1.00	1.00	.83	.71	.63	.56	.53

Proportion of Employees Considered Satisfactory =.60

Selection Ratio

r	*.05*	*.10*	*.20*	*.30*	*.40*	*.50*	*.60*	*.70*	*.80*	*.90*	*.95*
.00	.60	.60	.60	.60	.60	.60	.60	.60	.60	60	.60
.05	.64	.63	.63	.62	.62	.62	.61	.61	.61	.60	60
.10	.68	.67	.65	.64	.64	.63	.63	.62	.61	.61	.60
.15	.71	.70	.68	.67	.66	.65	.64	.63	.62	.61	..61
.20	.75	.73	.71	.69	.67	.66	.65	.64	.63	.62	.61
.25	.78	.76	.73	.71	.69	.68	.66	.65	.63	.62	.61
.30	.32	.79	.76	.73	.71	.69	.68	.66	.64	.62	.61
.35	.85	.82	.78	.75	.73	.71	.69	.67	.65	.63	.62
.40	.88	.85	.81	.78	.75	.73	.70	.68	.66	.63	.62
.45	.90	.87	.83	.80	.77	.74	.72	.69	.66	.64	.62
.50	.93	.90	.86	.32	.79	.76	.73	.70	.67	.64	.62
.55	.95	.92	.88	.84	.81	.78	.75	.71	.68	.64	.62
.60	.96	.94	.90	.87	.83	.80	.76	.73	.69	65	.63
.65	.98	.96	.92	.89	.85	.82	.78	.74	.70	65	.63
.70	.99	.97	.94	.91	.87	.84	.80	.75	.71	.66	.63
.75	.99	.99	.96	.93	.90	.86	.81	.77	.71	.66	.63
.80	1.00	.99	.98	.95	.92	.88	.83.	.78	.72	.66	.63
.85	1.00	1.00	.99	.97	.95	.91	.86	.80	.73	.66	.63
.90	1.00	1.00	1.00	.99	.97	.94	.88	.82	.74	.67	.63
.95	1.00	1.00	1.00	1.00	.99	.97	92	.84	.75	.67	.63
1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.86	.75	.67	.63

Proportion of Employees Considered Satisfactory =.70

Selection Ratio

r	*.05*	*.10*	*.20*	*.30*	*.40*	*.50*	*.60*	*.70*	*.80*	*.90*	*.95*
.00	.70	.70	.70	.70	.70	.70	.70	.70	.70	.70	.70
.05	.73	.73	.72	.72	.72	.71	.71	.71	.71	.70	.70
.10	.77	.76	.75	.74	.73	.73	.72	.72	.71	.71	.70
.15	.80	.79	.77	.76	.75	.74	.73	.73	.72	.71	.71
.20	.83	.81	.79	.78	.77	.76	.75	.74	.73	.71	.71
.25	.86	.84	.81	.80	.78	.77	.76	.75	.73	.72	.71
.30	.88	.86	.84	.82	.80	.78	.77	.75	.74	.72	.71
.35	.91	.89	.86	.83	.82	.80	.78	.76	.75	.73	.71
.40	.93	.91	.88	.85	.83	.81	.79	.77	.75	.73	.72
.45	.94	.93	.90	.87	.85	.83	.81	.78	.76.	.73	.72
.50	.96	.94	.91	.89	.87	.84	.82	.80	.77	.74	.72
.55	.97	.96	.93	.91	.88	.86	.83	.81	.78	.74	.72
.60	.98	.97	.95	.92	.90	.87	.85	.82	.79	.75	.73
.65	.99	.98	.96	.94	.92	.89	.86	.83	.80	.75	.73
.70	1.00	.99	.97	.96	.93	.91	.88	.84	.80	.76.	.73
.75	1.00	1.00	.98	.97	.95	.92	.89	.86	.81	.76	.73
.80	1.00	1.00	.99	.98	.97	.94	.91	.87	.82	.77	.73
.85	1.00	1.00	1.00	.99	.98	.96	.93	.89	.84	.77	.74
.90	1.00	1.00	1.00	1.00	.99	.98	.95	.91	.85	.78	.74
.95	1.00	1.00	1.00	1.00	1.00	.99	.98	.94	.86	.78	.74
1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.88	.78	.74

Proportion of Employees Considered Satisfactory =.80

Selection Ratio

r	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95
.00	.80	.80	.80	.80	.80	.80	.80	.80	.80	.80	.80
.05	.83	.82	.82	.82	.81	.81	.81	.81	.81	.80	.80
.10	.85	.85	.84	.83	.83	.82	.82	.81	.81	.81	.80
.15	.88	.87	.86	.85	.84	.83	.83	.82	.82	.81	.81
.20	.90	.89	.87	.86	.85	.84	.84	.83	.82	.81	.81
.25	.92	.91	.89	.87	.87	.86	.85	.84	.83	.82	.81
.30	.94	.92	.90	.88	.88	.87	.86	.84	.83	.82	.81
.35	.95	.94	.92	.89	.89	.89	.87	.85	.84	.82	.81
.40	.96	.95	.93	.90	.90	.89	.88	.86	.85	.83	.82
.45	.97	.96	.95	.93	.92	.90	.89	.87	.85	.83	.82
.50	.98	.97	.96	.94	.93	.91	.90	.88	.86	.84	.82
.55	.99	.98	.97	.95	.94	.92	.91	.89	.87	.84	.82
.60	.99	.99	.98	.96	.95	.94	.92	.90	.87	.84	.83
.65	1.00	.99	.98	.97	.96	.95	.93	.91	.88	.85	.83
.70	1.00	1.00	.99	.98	.97	.96	.94	.92	.89	.85	.83
.75	1.00	1.00	1.00	.99	.99	.98	.95	.93	.90	.86	.83
.80	1.00	1.00	1.00	1.00	.99	.99	.96	.94	.91	.87	.84
.85	1.00	1.00	1.00	1.00	1.00	.99	.98	.96	.92	.87	.84
.90	1.00	1.00	1.00	1.00	1.00	1.00	.99	.97	.94	.88	.84
.95	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.99	.96	.89	.84
1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.89	.84

Proportion of Employees Considered Satisfactory =.90

Selection Ratio

r	*.05*	*.10*	*.20*	*.30*	*.40*	*.50*	*.60*	*.70*	*.80*	*.90*	*.95*
.00	.90	.90	.90	.90	.90	.90	.90	.90	.90	.90	.90
.05	.92	.91	.91	.91	.91	.91	.91	.90	.90	.90	.90
.10	.93	.93	.92	.92	.92	.91	.91	.91	.91	.90	.90
.15	.95	.94	.93	.93	.92	.92	.92	.91	.91	.91	.90
.20	.96	.95	.94	.94	.93	.93	.92	.92	.91	.91	.90
.25	.97	.96	.95	.95	.94	.93	.93	.92	.92	.91	.91
.30	.98	.97	.96	.95	.95	.94	.94	.93	.92	.91	.91
.35	.98	.98	.97	.96	.95	.95	.94	.93	.93	.92	.91
.40	.99	.98	.98	.97	.96	.95	.95	.94	.93	.92	.91
.45	.99	.99	.98	.98	.97	.96	.95	.94	.93	.92	.91
.50	1.00	.99	.99	.98	.97	.97	.96	.95	.94	.92	.92
.55	1.00	1.00	.99	.99	.98	.97	.97	.96	.94	.93	.92
.60	1.00	1.00	.99	.99	.99	.98	.97	.96	.95	.93	.92
.65	1.00	1.00	1.00	1.00	.99	.98	.98	.97	.96	.94	.92
.70	1.00	1.00	1.00	1.00	.99	.99	.98	.97	.96	.94	.93
.75	1.00	1.00	1.00	1.00	1.00	.99	.99	.98	.97	.95	.93
.80	1.00	1.00	1.00	1.00	1.00	1.00	.99	.99	.97	.95	.93
.85	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.99	.98	.96	.94
.90	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.99	.97	.94
.95	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.98	.94
1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.95

Proportion of Employees Considered Satisfactory =.95

Selection Ratio

r	*.05*	*.10*	*.20*	*.30*	*.40*	*.50*	*.60*	*.70*	*.80*	*.90*	*.95*
.00	.95	.95	.95	.95	.95	.95	.95	.95	.95	.95	.95
.05	.96	.96	.96	.96	.95	.95	.95	.95	.95	.95	.95
.10	.97	.97	.96	.96	.96	.96	.96	.96	.95	.95	.95
.15	.98	.97	.97	.97	.96	.96	.96	.96	.96	.95	.95
.20	.98	.98	.97	.97	.97	.97	.96	.96	.96	.95	.95
.25	.99	.98	.98	[illegible]	.97	.97	.97	.96	.96	.96	.95
.30	.99	.99	.98	[illegible]	.98	.97	.97	.97	.96	.96	.95
.35	.99	.99	.99	.98	.98	.98	.97	.97	.97	.96	.96
.40	1.00	.99	.99	.99	.98	.98	.98	.97	.97	.96	.96
.45	1.00	1.00	.99	.99	.99	.98	.98	.98	.97	.96	.96
.50	1.00	1.00	1.00	.99	.99	.99	.98	.98	.97	.97	.96
.55	1.00	1.00	1.00	.99	.99	.99	.99	.98	.98	.97	.96
.60	1.00	1.00	1.00	1.00	.99	.99	.99	.99	.98	.97	.96
.65	1.00	1.00	1.00	1.00	1.00	1.00	.99	.99	.98	.97	.97
.70	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.99	.99	.98	.97
.75	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.99	.99	.98	.97
.80	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.99	.98	.97
.85	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.99	.98
.90	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.99	.98
.95	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.99
1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

APPENDIX U. REGRESSION MODEL WITH A UNIT VARIABLE

This appendix examines the final ATGP model with an additional variable for Unit. This variable is a binary varaible that may determine the effect of unit environment to recruiter success. The recruiters from the Baltimore Recruiting Battalion were assigned a value of one, and the recruiters from the Santa Anna Recruiting Battalion were assigned a value of zero.

Model with a Unit Variable

Regression Statistics	
Multiple R	0.33250312
R Square	0.11055832
Adjusted R Square	0.0637456
Standard Error	0.10992135
Observations	101

ANOVA

	df	SS	MS	F	Significance F
Regression	5	0.1427	0.0285	2.3617	0.0457
Residual	95	1.1479	0.0121		
Total	100	1.2905			

	Coefficients	Standard Error	t Stat	p-value	Lower 90%	Upper 90%
Intercept	0.51263259	0.0642	7.9895	0.0000	0.4061	0.6192
AFQT	0.00072462	0.0005	1.5069	0.1352	-0.0001	0.0015
TEST	0.00079547	0.0006	1.2494	0.2146	-0.0003	0.0019
GENDER	0.1166344	0.0575	2.0272	0.0454	0.0211	0.2122
PMOS	-0.02966972	0.0231	-1.2831	0.2026	-0.0681	0.0087
UNIT	0.0300826	0.0233	1.2891	0.2005	-0.0087	0.0688

From the statistics of the coeffiecients, it can be seen that the Unit variable is not significant at the 0.10 variable. A two-tailed test had to be used because it is uncertain which unit would be more successful. The affect of the Unit variable on the other variables are noteworthy. AFQT became more significant while Test became less significant, indicating that AFQT score distributions are different between the units, while SCT scores are relatively the same. PMOS also became less significant and Gender remained significant, even at the 0.05 level.

LIST OF REFERENCES

1. Fact sheet on "Review of Different Sales Aptitude Tests" provided by Professor Ronald A. Weitzman, Naval Postgraduate School, Monterey, CA, March 1994.

2. Bruce, M.M., *Examiner's Manual, Sales Comprehension Test*, Martin M. Bruce, Ph.D., Publishers, Revised, (undated).

3. McKenna, F.S., *Personnel Selection, A Self Instructional Program*, Educational Methods, Inc., (undated).

4. Matos, R.E., *US Navy's Delayed Entry Program: Effects of Its Length on DEP Loss and First Term Attrition*, Mast., Naval Postgraduate School, 1994.

5. Higman, R.P., *A Multiple Regression Model for Predicting Zone "A" Retention*, Mast., Naval Postgraduate School, 1986.

6. Crawford, A.B. and Burnham, P.S., *Forecasting College Achievement. A Survey of Aptitude Tests for Higher Education*, Yale University Press, 1973.

7. Wigdor, A.K. and Garner, W.R., *Ability Testing: Uses, Consequences, and Controversies*, National Academy Press, 1982.

8. Headquarters, United States Recruiting Command, *Smart Book*, December 1993.

9. Warnick, D. and Lininger, C.A., *The Sample Survey: Theory and Practice*, McGraw Hill Company, 1975.

10. Kearl, C.E. and Nelson, A., "*The Army's Delayed Entry Program,*" *Armed Forces and Society*, vol. 18, no. 2, pp. 253-268, Rutgers University, 1992.

11. Buchs, T.A., *Validation and Justification of the Use of a Sales-Aptitude Test for US Army Recruiter Selection*, Mast., Naval Postgraduate School, 1994.

12. Fact sheet on "Recruiter Reliefs" provided by USAREC Headquarters, PA&E Division, Fort Knox, Kentucky, January 1994.

13. Headquarters, Department of the Army, *Army Regulation 601-1, Assignment of Enlisted Personnel to the US Army Recruiting Command*, March 1992.

14. Eitelberg, M.J., *Manpower for Military Occupations*, Human Resources Research Organization, 1988.

15. Fact sheet on "Update of the Recruit the Recruiter Program' provided by USAREC Headquarters, PA&E Division, Fort Knox, KY, January 1994.

16. Weitzman, R.A., *The Prediction of College Achievement from the Scholastic Aptitude Test and the High School Record*, Ph.D., Naval Postgraduate School, 1981.

17. Klitgaard, R.E., *Choosing Elites*, Basic Books, Inc., 1985.

18. Fox, J., *Regression Diagnostics*, Sage Publications, 1991.

19. Berry, W.D. and Feldman, S., *Multiple Regression in Practice*, Sage Publications, 1985.

20. Headquarters, United States Army Logistics Management College, *Cost Estimating Reference Book*, October 1991.

21. Conover, W.J., *Practical Nonparametrics Statistics*, Second Edition, John Wiley and Sons, Inc., 1980.

22. Neter, J., Wasserman, W., Kutner, M.H., *Applied Linear Statistical Models, Third Edition*, Richrad D. Irwin, Inc., 1990.

23. Mendenhall, W., Wackerly, D.D., Scheaffer, R.L., *Mathematical Statistics with Applications*, PWS-KENT Publishing Company, 1990.

24. Fact sheet on "Recruiter Costs' provided by USAREC Headquarters, PA&E Division, Fort Knox, KY, May 1994.

25. Taylor, H.C. and Russell, J.T., "*The Realtionship of Validity Coefficients to the Practical Effectiveness of Tests in Selection: Discussion and Tables*" *Journal of Applied Psychology*, vol. 23, 1939.

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center 2
   Cameron Station
   Alexandria, VA 22304-6145

2. Library, Code 52 2
   Naval Postgraduate School
   Monterey, CA 93943-5002

3. Commander 1
   U.S. Army Recruiting Command
   Attn: PA&E Directorate (MAJ Poikenen)
   Fort Knox, KY 40121-2676

4. Director. 1
   U.S. Army TRADOC Analysis Command - Fort Leavenworth
   Attn: ATRC-FA (Mr. Graf)
   Fort Leavenworth, KS 66027

5. Headquarters, U.S. Army. 1
   DCSPER
   Attn: Colonel Durden
   Washington, D.C. 22301

6. Professor James G. Taylor. 1
   Department of Operations Research
   Naval Postgraduate School, Code OR/Tw

7. Professor Ronald A. Weitzman 1
   Department of Systems Management
   Naval Postgraduate School, Code SM/Wz

8. Major George C. Prueitt 1
   Department of Operations Research
   Naval Postgraduate School, Code OR/Pt

9. Captain Alejandro S. Hernandez 1
   DESA
   2251 Wyoming Blvd. SE
   Kirtland AFB, NM 87117